A Conceptual Framework for Monitoring the Emotional State of Students in VLEs

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Lubna Alharbi

March 2021

Abstract

The number of higher education students facing mental health issues has reached a critical point, with major consequences both in terms of academic achievement and general health and wellbeing. Institutions recognise the issue and have put in place a number of measures to try and counteract the crisis. Monitoring students' wellbeing can play a critical role, and Virtual Learning Environments could be instrumental in achieving this goal. This is particularly true in online-learning scenarios, where face-to-face contact is less present or not present at all, and educators need to rely on observing the online behaviour. However, the information overload to lecturers is significant, keeping track of each online discussion forum is extremely onerous, and the help from "learning analytics" not always useful, as they often concentrate on measures of students' performance, engagement, and presence in the virtual classroom, which are not necessarily good indicators of mental health.

The work presented in this thesis proposes to bridge this gap by addressing the effectiveness of emotion or writing style profiling to identify students at risk. The work proposes a conceptual framework for a system, intended to sit alongside the virtual learning environment, and able to play the role of an "emotion observer", identifying and flagging potential issues to the educator. We propose a system where technology is supportive of educators rather than replacing them, and is not intrusive or changing the classroom dynamics.

We demonstrate the validity of the approach with a series of experiments. We address the technical feasibility of such a system by investigating how established artificial intelligence techniques, and "off-the-shelf" tools implementing them, can be used to carry out the tasks that would need to be performed by a system implementing the approach, and we discuss their performances on either available datasets, or, for one of the experiments, a purpose built dataset, which is part of the contributions of the thesis. We address the admissibility of such an approach by conducting a focus group study with a group of experts in online learning.

The contribution of the thesis is therefore the first complete feasibility study on the development of a novel system able to monitor students' emotional state, both individually and as a cohort, and longitudinally over the course of their studies, which is aimed at supporting online educators identify students at risk and implement strategies for intervention.

Dedication

I dedicate this thesis

TO

my wonderful son Abdulaziz,
my beloved husband Zarea,
my lovely parents Sabriah and Abdulaziz

YOUR

love, help, support, patience, encouragements, and prayers are the strength that drives me in my entire life to complete this degree.

Acknowledgements

First and Foremost, thanks to Allah, "God", for his blessings of health, patience, and strength, empower me to complete my study. I am thankful to him for surrounding me with those people who support and encourage me during the most challenging and unforgettable experience ever.

I would like to express my sincere gratitude to my first supervisor Dr Floriana Grasso who has been given me the time, suggestions and guidance throughout the years under her supervision. Her research ideas, constructive criticism and valuable feedback, made the completion of my thesis possible. my sincere appreciation also go to my second supervisor Mr Phil Jimmieson; it is my privilege to work under his supervision. Many thanks for his invaluable support and advice.

I am also very grateful to my advisors Dr Valentina Tamma and Dr David Jackson, for their constructive feedback during the years of the research. My gratitude is also extended to Dr Mark Elshaw, my colleagues, and friends who provided me with valuable comments, support and assistance that enriched my personality and my work in many ways.

In addition, I owe a debt of gratitude to the Saudi Government, more specifically to Tabuk University for providing me with the opportunity and financial support that has enabled me to undertake my PhD.

Last but not least, I am deeply indebted to my husband and my son for standing beside me, you are the most important in my life without your love and support would impossible for me to achieve my goal. My great gratitude is also expressed to my parents and my sisters for their moral and emotional support.

Publications

The list of publications arising from this thesis:

1. L. Alharbi, F. Grasso, and P. Jimmieson, *A Conceptual Framework for Real-Time Emotional-State Monitoring of Students in VLEs to Identify Students at Risk* in *Human and Artificial Intelligence for the Society of the Future*, in Proceedings of the 2020 European Distance and E-Learning Network Annual Conference, Timisoara (and online), 22-24 June 2020. **(Note: This paper was shortlisted for the Young Scholar Award)**

2. L. Alharbi, F. Grasso, and P. Jimmieson, *A conceptual framework for identifying emotional factors leading to student disengagement in VLEs*, in Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education, ser. ITiCSE 2019. New York, NY, USA: Association for Computing Machinery, 2019, p. 293. [Online]. Available: https://doi.org/10.1145/3304221.3325570

3. L. Alharbi, F. Grasso, and P. Jimmieson, *An experiment with an off-the-shelf tool to identify emotions in students' self-reported accounts*, AISB'18 Symposium on Emotion Modelling and Detection in Social Media and Online Interaction, 2018.

4. L. Alharbi, F. Grasso, and P. Jimmieson, *Emotions detection to investigate lecturer-student relationships in the electronic learning environment*, 24th Annual Conference of the Association for Learning Technology, 5-7 September 2017.

Contents

List of Figures

List of Tables

List of Abbreviations

VLE	Virtual Learning Environment
LMS	Learning Management System
LA	Learning Analytics
AC	Affective Computing
NLP	Natural Language Processing
DM	Data Mining
NBC	Naive Bayes Classifier
ANN	Artificial Neural Network
SVM	Support Vector Machines
POS	Part of Speech Tagging
TP	True Positives
TN	True Negatives
FP	False Positives
FN	False Negatives
FGDs	Focus Group Discussion
MOOCs	Massive Open Online Courses

Chapter 1

Introduction

1.1 Motivation for the Work

Higher education is witnessing a mental health crisis. The largest mental health survey by the institute for Public Policy Research [200] reported that the number of students who have been diagnosed with a mental health condition has risen dramatically over the past 10 years, as well as the number of students at risk of dropping out because of the lack of support and care [19].

Similar findings were published by HESA, the Higher Education Statistics Agency[1], who reported that from a survey of 38,000 UK students, 21.5% had a mental health diagnosis, and 33.9% has experienced serious mental issues for which they felt they needed professional help, with the levels of anxiety, loneliness and thoughts of self-harm on the increase. Around 15,000 first year students in UK Universities reported a mental health condition in 2015/2016, and students in their second and third years are significantly at higher risk than students in their first year. There are many factors contributing to this problem, pressures from work, which is worsened by financial concerns, moving away from home, but also strict regulations

[1] `https://www.hesa.ac.uk`

and academic requirements of the educational institutions have been shown to provide a pressured environment for students, putting them more at risk to become stressed [102]. The impact of this is very serious, both in terms of the academic progress, and long term health effect, or in the worst cases, self harm and suicide.

Universities UK, an organism collecting 137 universities in England, Scotland, Wales and Northern Ireland, has long recognised the urgency of the issue [205] and published a number of reports and recommendations, the latest of which [206], in May 2020, recommends a *whole university approach*, where *"all aspects of university life promote and support student and staff mental health"*, from the learning environments, to the support services, to the staff mental health, to the environment.

The approach is confirmed by previous research, for instance, on the importance for both students and staff to be aware and the teachers and students are aware of the relationship between working environment and stress [30, 192, 193].

Mental health issues are not only confined to students moving to campus, away from home: online students are often forgotten in the equation, but the rising amount of students using this mode of learning has started to raise awareness of mental health issues among this type of students too [18]. Students learning online might have different needs, for example having different socio-economic backgrounds, need to juggle work and family [40], which puts them at higher risk, as well as encountering a higher sense of isolation and lack of social relationships with other students. The issue has become all too evident with the recent massive move to online learning as a consequence of the lockdown measures imposed during the COVID-19 pandemic [114, 173].

1.2 Problem Statement and Scope

The research presented in this thesis, is motivated by the above scenario, and intends to address the issue of monitoring the mental wellbeing of students working on a Virtual Learning Environment (VLE) [213].

VLEs are of course the foundation technology for modern academic studies, and provide the bonus of creating a rich environment to generate data for analysis. Higher education institutions already exploit the fact that this data can inform learning strategies, develop new areas of educational research and improve the quality of the overall organisation (including finances) [44, 183].

The analysis of the data generated from students' interactions in educational technologies underpins much of the computing educational research aiming to enhance or personalise the student learning experience. Usually, students communicate with the lecturer asking questions, complaining, or seeking pieces of advice. This happens overtime, during the course of an entire module, and possibly a programme of study. The hypothesis underpinning this thesis is that the detecting mental states, and in particular emotions, in these exchanged messages might help to identify any mental issues, and provide institutions a mechanism to give support at the proper time. This would support the development of an emotional profile for students which could in turn help identify not only whether the student is manifesting a specific distress or emotional engagement issue at a particular moment in time, but also monitor the situation over time, and identify changes in patterns.

In particular, we concentrate in this thesis on a scenario in which communication happens in an online, text-based classroom: while advanced technology provides the opportunity to use keyboards, cameras, and microphones to detect emotional states with a physical manifestation [106] we want to focus on a practical solution that can be applied very broadly, and not confined to very special cases in which cutting edge technology is available to both students and teachers.

A further hypothesis underpinning our research, we concentrate on a scenario in which technology is of support to educators. As prime contact for students, lecturers plays a crucial role in helping students not only on the academic side, but often in supporting them for a vast variety of issues students face. Working on VLE environments, where students regularly interact with their peers or discuss course content or performance, student monitoring becomes a challenging and demanding task, which adds to their educator role.

Therefore, in the current study, we want to explore the extent to which Artificial Intelligence technologies can assist lecturers, rather than replacing them, in addressing the issue of student mental health monitoring, by means of tools that can sit alongside a VLE and equip the VLE with the capability to detect changes in the emotional state, with a view to flag situations for the lecturer to intervene at proper time.

The scope of the work, while exploring and analysing AI technology, sits within the area of Computers in Education, and will explore how the combination of various areas, Text Analysis (classification by sentiment/emotion and writing style features), Emotion Modelling and e-Learning/Learning Analytics, can provide a conceptual framework able to address the problem defined.

1.3 Research Questions

The following research question is therefore extracted from the problem statement above:

> **RESEARCH QUESTION:** ***Can we model a real-time system, incorporated in a Virtual Learning Environment (VLE), able to observe the emotional state of a student or cohort of students longitudinally, and use this model to help an e-learning mentor make targeted time-***

sensitive interventions?

This question can be addressed by investigating the following sub-questions:

RQ1: To what extent textual messages students exchange in a VLE can be reliably analysed in terms of their emotional load?

RQ2: Can we model a system able to observe the behaviour of a student or group of students in a VLE overtime and create an emotional profile, in order to identify changes and flag outliers?

RQ3: How confident can we be that such system can be both technically feasible and acceptable for uptake by educational stakeholders?

1.4 Overview of the Thesis

This thesis will address the previous research questions by conducting a series of experiments, and will be structured as follows:

1.4.1 Literature Review and Background Research

In ***Chapter[2]*** we will explore previous research and required background knowledge. The literature review is divided into four main parts: I) Virtual Learning Environment, II) Learning Analytics, III) Sentiment and Emotion analysis, IV) Writing Style features.

1.4.2 Framework and Methodology

In ***Chapter[3]*** we introduce our pedagogical scenario and propose a conceptual framework to address the research question. We then discuss the methodology used to evaluate such framework, introducing the set of experiments conducted.

1.4.3 Experiments

The experiments are discussed in three Chapters, each of them including an overview and purpose of the experiment, the methodology used, and a discussion on the results and limitations. The experiment in ***Chapter[4]*** will squarely address RQ1, the one in ***Chapter[5]*** will follow up by dovetailing RQ1 with RQ2 and discussing the ability to use emotion identification to build a student profile able to make predictions on student performance, while ***Chapter[6]*** will address RQ2 and discuss how a longitudinal emotional profile of students and cohorts can be created. In the three experiments, we rely on existing general datasets. In the first experiment we use a dataset collected in an educational situation setting, though not captured during learning activities, consisting of self-reported emotional related sentences. In the second experiment the dataset comes from actual class interactions from the Stanford MOOCs collections of datasets. In the third experiment, the dataset is not related to education, but was adapted from a Motivational Interviewing corpus. In the Chapter related to each experiment we discuss the benefits and shortcomings of each of these choices.

1.4.4 Evaluation

In ***Chapter[7]*** we move from technical feasibility to acceptability evaluation, therefore addressing RQ3, and find out whether practitioners in the field would find our approach admissible and feasible.

1.4.5 Conclusion

Finally, ***Chapter[8]*** will conclude the thesis with a reflection on achievements, limitations, and further research directions.

Chapter 2

Related Work

The previous chapter provided an understanding of the personal interest and inspiration that lead to the undertaking of this research. The chapter also provided the background and outline of the research study. This chapter will review some of the previous work and other attempts done to solve the research problem. It will illustrate how our research work belongs to the area of Emotional Intelligence and VLEs. Also explored in this chapter is a review of the Learning Analytics (LA) and its application to VLEs. This is done in order to recognise some key areas for further investigation.

2.1 Virtual Learning Environment

2.1.1 Overview

A Virtual Learning Environment (VLE)- sometimes known as a Learning Management System (LMS)- is a computer-based environment usually within educational institutions, designed to provide a personalised education by means of accessing e-learning courses effectively anytime and anywhere [27, 213]. Users of a VLE are usually divided into two major classes: students and tutors [27]. Several studies report that in comparison with a traditional learning institution, a VLE is more productive, more effective, more satisfying and is often used as a place to collaborate and extend discussions between instructor and learner [101, 127, 159].

The explanation of why VLEs have become so common and integrated into many higher education institutions is that the VLE provides support, resources and good communication links [101]. In addition, a VLE provides real benefits for institutions to increase the number of students, and in doing so supports widening participation. Increasing numbers of students on campus is a clear problem for institutions [199], which a VLE could help alleviate. In terms of widening participation, a VLE can provide support and resources for part-time students who are not always able to travel to the campus.

A VLE can facilitate the same teaching and learning principle features (e.g. teacher-student communication, peer support, group work, self-assessment, tutorials) that traditional face-to-face teaching can offer. A VLE manages information related to student teaching. For instance it can support online discussions, lecture notes, and the calculation and reporting of student grades. Dimensions and examples of principal components of a VLE are provided in Figure [2.1] to show how a VLE differs from traditional classrooms [27]. *Storage and Distribution of Learning Materials, Assessment and Feedback Tools, Collaborative Activities, Progress Tracking and Monitoring* are the main features of VLEs.

Storage and distribution of learning materials is a principal feature of VLEs, enabling the delivery of content in a variety of formats. Lecture support notes are loaded in PowerPoint presentations, images, audio or video. An advantage of this is that the tutor can deliver the content in a way to achieve the appropriate learning goals. Evaluation and feedback tools are a key feature provided by VLEs to support the assessment of students online. VLEs provide several different types of assessment, such as self-tests for formative feedback and

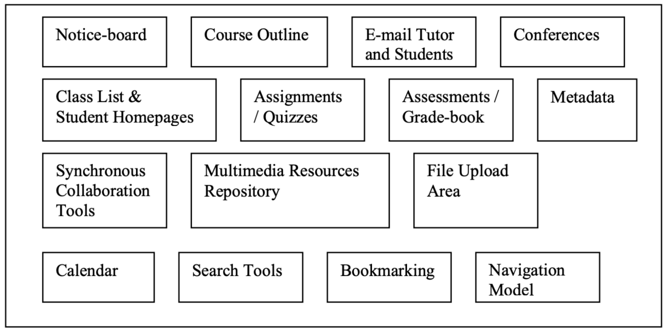


Figure 2.1: A schematic of a prototypical VLE adapted from [27]

automated online quizzes.

VLEs allow tutors to provide constructive and meaningful feedback for various tasks. Collaborative activities among students enable them to discuss new ideas and share information. Many tools (e.g. Live Chat, Shared Whiteboards, and Video Conferencing) are used to support collaboration within and across student groups. Progress tracking, and monitoring are a powerful components of VLEs and have a wide range of benefits for both students and tutors. They can provide tutors with information about the first time a student has accessed a course, how often they accessed it, and in what areas.

There is a range of different VLEs available, which all have the same functionality but work slightly differently. These include Blackboard, WebCT, Lotus LearningSpace, and Moodle. According to the Times Higher Education Magazine, Blackboard is widely used in higher education [211]. The Blackboard Learning System allows teachers to post class details and course materials, lectures and assignments, and provides essential communication and other interactive resources [195].

2.1.2 E-learning Systems in a Virtual Learning Environment

E-learning is a learning system that is delivered via the internet and which features an interactive mode to support teaching and learning. Over recent years there has been a massive increase in the use of e-learning courses as a way to complement traditional approaches to teaching [79, 101]. E-learning is one of the modern methods available to deliver education in universities. There are a number of students studying via e-learning courses in public and private organisations who have no regular face-to-face contact with a teacher in the classroom [170].

A comparison between e-learning and the traditional classroom shown that e-learning has its advantages in student learning outcomes [111]. Ensuring that all students participate in face-to-face sessions is extremely difficult, but e-learning provides a mechanism for all students to contribute. E-learning techniques can offer individualised content; a learner could access the material anytime, uploading their homework, doing quizzes and submitting assessments [27, 187].

At the same time, an e-learning environment allows teachers to support students' learning outside of lectures at anytime as e-learning techniques reduce the need for face-to-face contact time with students. Teachers track student performance, upload information, create or modify the content at a time and place that suits their schedules. Teachers can provide additional reading materials to students to help them understand the concepts and ideas in sufficient detail.

One of the most significant features of e-learning courses is the communication and collaboration between students and their peers, or between students and their teachers. This is facilitated using discussion forums, wikis, and blogs. Online discussions can be used to raise discussions about any problems or to ask questions; teachers could participate by providing appropriate feedback. Another significant feature is that teachers can embed links to relevant videos from YouTube or to other sites.

A valuable aspect of the e-learning environment is its ability to provide detailed data about learners' engagement with materials and activities. This can help support students who are struggling and flag up issues earlier than might otherwise be possible. E-learning teaching has some difficulties, yet it is still a powerful and promising educational strategy [146]. Loneliness and feelings of isolation are known to be a potential negative aspect of e-learning. As a result of this most students engaged in e-learning at some point struggle with their course, or feel like disengaging from it [146, 170, 223].

Pankaj Srivastava [191] stated that student isolation could lead to problems with engagement and motivation, as it is one of the main barriers to e-learning when compared with traditional classes. Another recognised issue in the e-learning environment is that teachers could be disconnected from the needs and interests of the students [191, 223]. Some student might feel lost when they cannot ask an instructor for more clarifications regarding the material. A potential risk of e-learning courses is the lack of motivation, especially for students who struggle to study independently [158]. In addition to this, some students do not like the use of online tools and do not contribute in group discussions.

2.1.3 Early Disengagement Investigation in the E-Learning Environment

There is no official definition of engagement; generally, engagement means the interactions between the environment and the individuals. In education, engagement measures students' perceptions, efforts is a good indicator of their performance and achievements [62, 217]. The engagement level is different depending on the learning environment context. However, Fei et al. [111] argued in their study that there are no significant differences in students' engagements if there is a change in their learning environment from the traditional classroom to e-learning.

Research has repeatedly demonstrated that student disengagement has been a concern for a variety of reasons. Student motivation and engagement are firmly related components of student learning that can affect learning results. Engaged students could enhance key performance factors such as grades, attendance, student satisfaction and University reputation. Vicki Trowler [202] states that, despite the fact that there is no generally acknowledged meaning of what constitutes "engagement", student and school achievement, student maintenance and student motivation are constantly connected to engagement. For instance, a portion of the early studies characterised engagement in a number of different ways; for example, interest [49], exertion [128], time spent [20] and motivation [185].

One way to improve the learning process and students' progress is to monitor each student in their online activity and recognise the main problems that the students face [32]. A study conducted by Docherty [51] proposed Computer Supported Learning (CSL), which is a student-centered, problem-based approach to the gaining of education and knowledge. They used web pages, rather than traditional lectures delivered by lecturers. The idea of using a specialised environment to fit the capabilities of each student was presented in [51].

The evaluations were positive (students were satisfied, benefited from improved performance in exams, and increased their self-efficacy). These results indicated that using information technology that adjusts the educational material to fit the students' capabilities may satisfy students and encourage them to perform well. However, Docherty [51] removed the role of lecturers and relied totally on the system to educate students. It may be argued that it would be better to involve the lecturer in this process to monitor and advise the system [57] enabling better preparation of educational material, tailored to specific students after a review of their performance.

2.2 Learning Analytics

2.2.1 Learning Analytics Overview

Learning Analytics (LA) means "the measurement, collection, analysis and reporting of data on students and their contexts, to understand and adjust learning and the environments in which it occurs" [183]. According to Brown [29], learning analytics (LA) means the process of improving learning through systematically collecting and analysing large data sets from online sources. Big data sets could be from different data sources such as open-source platforms (e.g., Moodle), learning management systems (e.g., Blackboard), and open social platforms (e.g., LinkedIn) [164].

Educational institutions that are using virtual learning environments present an ideal context for the use of LA as they already have a very significant amount of data available for analytics research with their large numbers of students and the growing use of the internet and mobile technology. Moreover, LA focuses on building systems that are able to adjust the content, support levels and other customised services by constantly gathering, recording, capturing, and acting on data in a way that minimises the time delay between data capturing and use of these data [16].

Therefore, unlike existing assessment processes that use the outcomes of one semester to guide changes in the next semester, learning analytics aims to combine historical and current user data to predict what services individual users may find useful now. Thus, learning analytics aims to build on the modelling capacity of analytics: predicting behaviour, acting on predictions and then returning those findings to the system to strengthen predictions over time as they apply to teaching and learning practices [56].

Gašević et al [68] summarised three common themes in the implementation of the LA, predictors and indicators, visualisations, and interventions [29]. The first category includes the use of statistical and data mining methods where the data collected from different factors (e.g., academic performance, students remaining in a course, student engagement) is used as a prediction model. These are generically known as Early Warning Systems [107] or those detecting students at risk [11].

The second category is to help students to interpret the information correctly from visualisations or dashboards, as complex data visualisations or dashboards are not recommended [68]. The third category focuses on interventions which generally involve how to shape the learning environment with precise actions to enhance student experience [218]. There have been several efforts from educational institutions to assess student learning by processing students' data to improve institutional decision-making and development [219].

Empirical studies have shown that LA may be useful for improving learning. LA raise awareness of learners and educators in their current situations, so they can make constructive decisions and perform their tasks more effectively [175]. Students' interaction could be increased as more of their activities are measured through the system. The teacher can take prompt follow-up action according to students' analytics level of interactions. Students who got support and personal interventions from their lecturers showed better academic performance, retention and had higher graduation rates [165].

Different universities have applied LA in their institutions [61, 219]. Albany Technical College and Ball State University are examples of using LA to improve efficiency by reducing the time required to diagnose issues and address specific problems with counselling. Moreover, they help students to consider the differences between their activities and expected outcomes, and the elements of their academic success [92, 220]. Oxford Brookes University is another example of using LA by monitoring students to improve their experience and to evaluate the modules and programs [179].

A study by Rienties [165] for the Open University (UK) used LA to deliver personalised intervention for students to achieve cost-effectiveness. Other universities focus on monitoring student attendance, such as The University of East London. In that institution, LA data was used to trigger the sending of e-mail warnings to students without satisfactory attendance [179]. Hanover Research defined different types of analysis, both for learning analytics and institutional analytics [76]. Table [2.1] lists several typical implementations identified in the literature and in case studies by major category (learning vs. institutional) and analysis level [13, 22, 37, 76, 177].

ANALYTICS CATEGORY	LEVEL OF ANALYSIS	TYPE OF ANALYSIS
Learning Analytics	**Course Level**	- Social Network Analysis - Conceptual Development Analysis - Discourse Analysis - Personalized Curriculum - Student Performance Assessment - Degree Audit
	Student Level	- Performance Assessment - Predictive Performance Analysis / Early Warning Systems - Automated Advising and Coaching
	Departmental Level	- Early Warning / Predictive Modelling
Institutional Analytics	**Instructor Level**	- Teacher Effectiveness - Financial Contributions
	Student or Student Body Level	- Enrolment Profiling and Predictive Analysis - Lifetime Value / Booster Effectiveness - Advocacy - Post - Educational Employment Analysis - Subject or Course Selection Recommendations
	Institutional Level	- Admissions Analysis - Institutional Performance/Efficiency - Retention/Attrition Trends
	Public Level	- Comparison with Other Institutions

Table 2.1: Examples of learning and institutional analytics employed in Higher Education, adapted from Hanover Research [76]

LA could be used as a way to identify factors that have significant statistical correlations with final course outcomes. A research study by Grush [73] monitored the online students learning at Rio Salado College by using the Progress and Course Engagement (PACE) system to automatically track student progress with interventions as needed. Teachers can easily see who is at risk after the first week of the course as the system generates a report with green, yellow, and red flags. Results showed 70 percent accuracy as to whether any given the student would complete the course successfully using three factors: read the course materials online, do practice exercises, and how many points they are getting on their assignments [73].

All these examples show how educational stakeholders (e.g., learners, instructors, and administrators) could benefit from these analyses embedded in applications to map students behaviours, the duration, relative effort, and trends within better performance including academic performance level. Bench-marking is built in as this can be compared to other students within the same categories or the same module. It helps to monitor and focus on key areas to improve and reflect on performance during activities. Likewise, it will flag up dips in effort, performance and lessening commitment.

The University of Maryland Baltimore County (UMBC) found that students earning below a C showed continuously 40 % lower use of the LMS compared to students earning a C or higher [65]. UMBC launched the "Check My Activity (CMA)" feedback tool, which provides students with their performance and compares students' activity to a summary of anatomised peer data. Students could see an outline of their performance in all Blackboard subjects compared to the average of the class. The information is collected from the number of times the student has signed into the session and the number of times the student has viewed a file or posted to the discussion board.

2.2.2 Applying Learning Analytics to Predict Students' Engagement

In higher education, student engagement has been described as an ongoing issue, and learning analytics may play a central role in retention efforts. Pekrun [153] considered engagement as one of the qualitative measures of the learning process, and a possible predictor of students' academic functionality. Learning Analytics is a data analysis method that can provide insights into student behaviour to enhance academic performance outcomes [186].

Several uses of learning analytics include systems for early warning/risk identification for student retention, course selection recommendations, and student progress tracking dashboards. Detecting students' engagement at an early stage requires a tracking system and advanced analysis of data. One of the key applications of LA is to track and predict the performance of learners and to identify potential problems and provide an early warning for students at risk [11, 16].

Experts claim that data analytics will be commonly used in online education over the next few years to recognise the behaviour patterns of students and increase the learning and retention rates of students [16]. The benefits include processing the data from a student's interaction with their online environment to boost learning and teaching. Another important utility of LA is used as a way to improve student maintenance.

For example, highlighting where supplementary support is required at an early stage [11, 50]. Purdue University provides preemptive support to students who struggle in subjects that use their data gathered from Blackboard. The instructor observed students in four major metrics components: subject performance (e.g. Points earned to date), effort (e.g. Interaction with LMS compared to peers), prior history (e.g. Academic preparation, GPA, test scores), and student characteristics (e.g. Residency, Age) [12, 160].

LA presents the use of timely feedback and intervention. Teachers may receive up-to-date holistic information about students' study progress. They could identify students who are isolated, performing poorly or behaving differently from others [198]. Higher education institutions are now using statistical analyses of students to determine better which students are at higher risk of dropping out or failing to complete their study course, and use the data statistics to intervene before this occurs.

The University of Michigan [126] used LA to identify at-risk students and provide personalised feedback to each student. The University of New England captures the student learning well-being status to provide timely support for students studying part-time and encourage peer-to-peer student networking [179].

In previous research, researchers use a student's educational history and demographic information to predict the student's performance. A research study was carried out by Marbout [124], to predict at-risk students as early as possible. They used different machine learning algorithms such as Naive Bayes Classifier (NBC), Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), in order to process some input features, such as grades, attendance, and quizzes.

Thus, a study by Hussain [82] uses machine learning algorithms in specific educational analyses, such as average time, the total number of activities, average idle time, the average number of keystrokes and overall related activity for different exercises during learning sessions in order to predict student difficulties in advanced exercises. So, before the start of the upcoming session the instructors of an e-learning system can warn students in advance of the difficulties they will encounter.

Different sources of data in the learning management system could be analysed to understand students' learning engagements. For instance, using LA helps capturing the attitudes of the students while watching videos of the course, by showing the time spent watching videos or the most frequently watched parts of the videos. The number of assignment uploads and course views are used by Sisovic [184] as predictors of academic success on programming courses. Research in [39] found that students who visit their pages during the first week of the semester achieve a higher final grades result.

Dietz [50] examined the use of learning analytics in educational institutions and how teachers can use effective tools for learning analytics to monitor and estimate student performance, thereby predicting whether students will fail or pass.

A student might be informed of their progress through different date, including: a set of subjects, number of hours spent in a course, number of 'streaks' or days that they get into online content, their quizzes' performance or comparison of the student's activity with their peers. These analytics can be used to increase commitment and motivation. The illustration on how students may obtain their own educational analytics can be highly motivational and can help them focus on enhancing results. Therefore, we need to be mindful of the potential negative impact of student-related analytics which could have an opposite effect and create real angst in some students.

2.2.3 Gaps in Current Learning Analytics Prediction Systems

As discussed above, learning analytics systems are a growing area of interest that are commonly associated with attempts to monitor and explore a large amount of data produced in virtual learning environments. Most VLEs have built-in tracking tools which provide information on each student. LA aims to process students' data in order to improve the student's performance, including measures to enhance participation, increase the completion of the course, and minimise the study time required before course completion and successful graduation [121].

Many LA use dashboards to provide visualisations that display statistical information such as time spent online, logging and activity. This enables teachers and learners to make informed decisions about the learning process. The literature demonstrates that visualisation would be a valuable tool for education [85]. The visual tools must be designed to offer a sufficient interpretation for the tasks they are supposed to support [207].

Visualisations of learning analytics may be harmful if not carefully designed and used. A recent literature survey found that the majority of dashboard visualisations could be more useful for teachers than for learner [100]. Some learners reported that seeing their own engagement or comparing them with their peers would not positively support them to perform well [86].

For VLEs, where students regularly interact with their peers or discuss course content or performance, a challenging and demanding task for the e-teachers is the assessment and tracking of students' academic performance. Most LA are looking for data statistics only and ignore the context of what the actual conversations are about. Researchers claim that tracking functionality in most VLEs provides teachers with inadequate learning activity reports which do not meet their students' needs effectively [56, 79, 226].

Several studies reported that students who perform most of their activities and spend most of their time online would not be highly efficient in their learning and would not have the highest academic performance [103, 117, 118]. On the other hand, there is a common assumption, for instance, that time spent on learning is positively related to academic performance [64].

A more thorough study of learning analytics involves the development of the new direction of including emotions as an early warning to identify students at risk. This is the case at the University of New England (UNE) [179], unlike what happens in the other institutions discussed in the previous section. UNE introduced an early warning system to classify at-risk students by incorporating subjective and emotional data into its early-warning analytics.

Students represent their feedback about their current feelings about the subject. They recorded how they feel using a free text response box as well as a set of emoticons to choose from (happy, neutral, unhappy and very unhappy). The Student Support Team contacts any students with negative emotions within 24 hours. This study showed the effective use of qualitative feedback from students to increase students' motivation [179].

In this thesis, we argue that few studies have paid sufficient attention to the analysis of the context, for example by using emotion and/or sentiment analysis in addition to the student's educational history for better understanding and prediction of student engagement in advance. We use real-time models to mix those techniques for maturing educational platforms in order to provide for reliable learning in real-time. At-risk student identification must be conducted as early as possible to allow sufficient time for instructors to conduct timely educational interventions to facilitate students' learning achievements.

2.3 Emotions and E-Learning Systems

2.3.1 Emotion and Related Concepts

Emotions are defined as the mental status associated with a wide variety of feelings, thoughts, and behaviours [31, 42]. According to David [144], the emotions of humans includes "...physiological arousal, expressive behaviours, and conscious experience.". Other concepts are used as well, intertwined with emotion are affect, feeling, sentiment, opinion, view, belief and cognition.

Affect has to do with value. It is a superordinate concept that subsumes particular valenced conditions such as emotions, moods, feelings and preferences [148]. The primary difference between emotions and feelings is that feelings are often experienced consciously while emotions tend to manifest themselves either consciously or subconsciously and are derived from a person's moods. For instance, a person may feel bad about a particular situation but not be emotionally affected, which makes it easier to cope with the situation as they were not emotionally attached to the situation. Feelings are learned behaviours that are often in hibernation until triggered by external events. Emotions are therefore event-driven [143, 212].

Sentiment, on the other hand, is an individual's particular opinion or view. Opinions are beliefs that people tend to form about a particular issue which determines their reaction to such issues. The views of a person can be said to be their beliefs. Beliefs are therefore thoughts about a particular situation or event and can be changed through learning. Beliefs may be rationally based on learned facts or irrationally based on heard rumours. Cognition, on the other hand, relates to the mental functions dealing with logic as opposed to affective dealing with emotions [143, 212].

2.3.2 Affect Theories and Models of Emotions

Affect theory is a philosophy that attempts to classify affects into discrete categories and used interchangeably with emotions, or feelings. Researchers have been studying emotions since Darwin [43] and many theories have been produced by different psychology schools that represent ways of understanding affective states. The primary emotional hypotheses can be divided into three main categories: physiological, neurological, and cognitive.

James-Lange Theory of Emotion [84], Cannon-Bard Theory of Emotion [33], and Silvan Tomkins' Affect Theory [201] are examples of Physiological theories which suggest the human body is the controller of emotions. Neurological theories (e.g. Facial-Feedback Theory of Emotion) argued that emotional reactions are a result of action within the brain [74, 93]. Lastly, cognitive theories (Cognitive Appraisal Theory [108]) proposed that feelings and other mental activities play a key role in formalising emotions.

Tomkins [201], defined "affect" and introduced the concept of the basic emotions and gave two labels for each emotion: one for low intensity and one for high intensity. When analysing eight emotions (excitement, joy, surprise, distress, disgust, anger, shame, and fear), he claimed that there are various circumstances that could raise one's affect or reaction. Tomkins organised each affect with its typical response, for instance, positive emotions enjoyment/joy is a reaction to success/impulse to share, and the responses are smiling, lips wide and out, or neutral, where, for instance, surprise/startle is a reaction to sudden change/resets impulses which is represented by eyebrows up, eyes blinking. A negative emotion, for instance, anger/rage is a reaction to threat/impulse to attack and represented by frowning, a clenched jaw, a red face [201].

Several emotional models were created, in which categorical and dimensional are considered as the primary two representation paradigms for emotions [188]. Affective researchers in computing use either a categorical model or a dimension model to classify emotions. Findings show that a dimension model depends upon the set of emotional categories [188]. The categorical theory proposes that a number of limited emotion categories are defined. Ekman [55] considered six basic emotions: anger, disgust, fear, happiness, sadness, and surprise, to which all others can be reduced. Plutchik [161] postulated a new conception of emotions in 1980, which he called "wheel of emotions" Figure [2.2].

This concept demonstrates how different emotions could merge with one another and generate new emotions. Plutchik [161] first proposed eight primary bipolar emotions: fear versus anger; sadness versus joy; disgust versus trust; and anticipation versus surprise. Similarly, the theory of Parrots [152] identified over 100+ emotions and conceptualised them as a tree structured list, where the first level is composed of six primary emotions (love, joy, surprise, anger, sadness, fear) Figure [2.3].

Russell [171] developed the two dimensions of emotions valence (negative/positive) and arousal or activation (low/high). Russell's dimensional model is named "the circumplex model of emotions". It contains eight types of emotions and it is possible to place every single emotion on this two dimensional graphic; the y-axis is the degree of arousal and the x-axis is the valence Figure [2.4] [87]. For instance, fearful is a negative emotion that creates high activation levels; relaxed is a positive emotion with relatively low activation.

The diagram shown in Figure [2.5] illustrates Lovheim "cube of emotion" model, which is based on Tomkin's eight basic fundamental emotions [201] (excitement, joy, surprise,

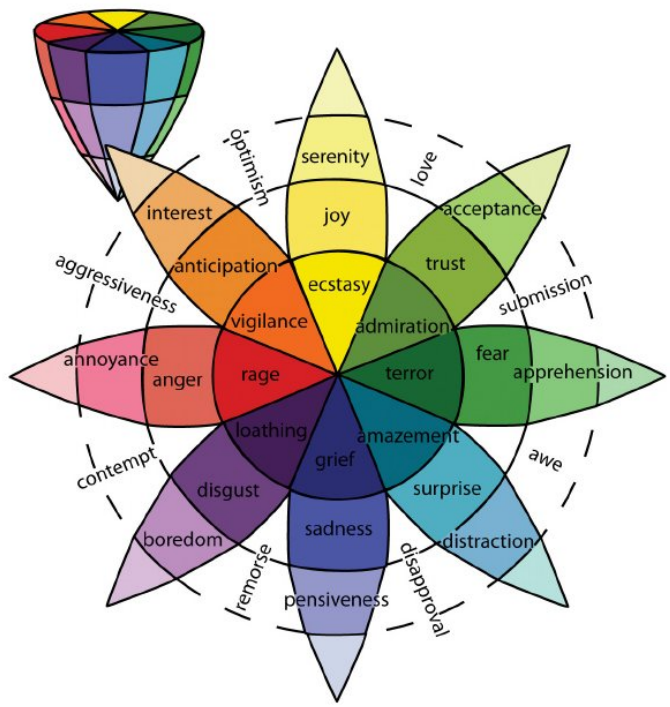


Figure 2.2: Plutchik's wheel of emotions and relationships between basic and derivative emotions [161]

distress, disgust, anger, shame, and fear). It was developed in 2012 as a theoretical model that focuses on central brain monoamines neurotransmitter interactions and the emotions that we experience [115]. Every corner of the cube has two labels of emotion, both labels represent high and low emotional intensities. The extreme and maximum emotional intensity are represented by the corner, while the lower intensity (a neutral state) is located between the corner and the centre of the cube.

Using these various computational models, the emotion in the text that evokes the emotional state of the writer can be interpreted. In this thesis, emotions are detected using tools that depended on categorical models. These tools will help to convey a particular set of features of emotions to assess our argument. The majority of research in affective

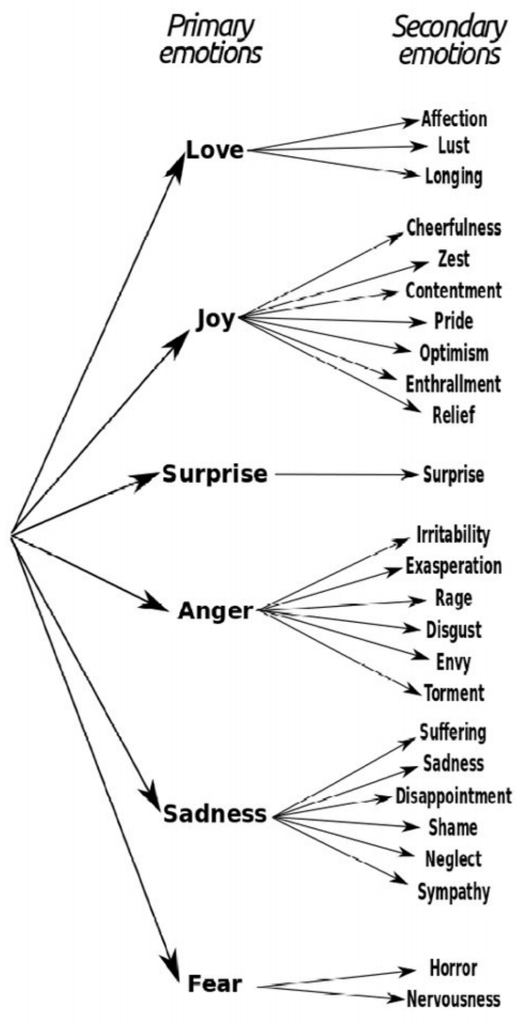


Figure 2.3: A tree-structured list was proposed by Parrot [152]

computing has focused on the six basic emotions listed above [190].

2.3.3 The Role of Emotions in Learning

Measuring students' emotional engagement is critical in the students' learning processes. Chaffar and Frasson [36] argued that the reason behind the failure to achieve efficient learning is mainly because of a lack of emotional intelligence abilities. Students' positive

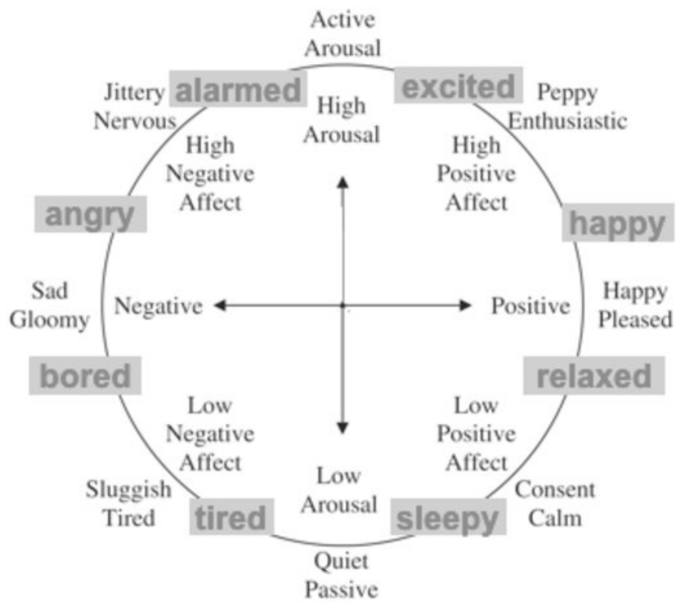


Figure 2.4: A dimensional model proposed by Russell [171] adapted from [87]

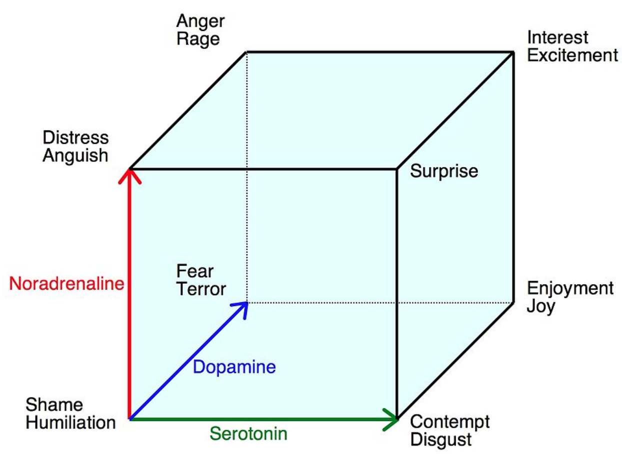


Figure 2.5: Lovheim's cube of emotion

and negative emotions may differ depending on the e-learning context [156]. For instance, students working on their dissertation, a long and hard task involving more grit than typical class assignments, might have very different educational and emotional needs. Depending on the task, students might feel overwhelmed by the specific information they found or might feel frustrated about the particular result they cannot achieve. Mega [129] found that students' positive emotions enhance their beliefs on the incremental theory of intelligence and confidence in their intelligence.

Pekrun [156] has created the academic emotion description which includes both affective states (such as confusion, anger, stress) and cognitive states (e.g. interest, boredom, fatigue). This study established classes of emotions and showed that positive emotions could predict creative thinking, foster good educational results, whereas negative emotions are more likely to be correlated with low grades. It was also suggested by Kim [96] that while learning, negative emotions can change the use of cognitive strategies and motivation.

Implicit or explicit information about students positive or negative emotions could influence the teacher to enhance the learning experience of the students and keep them from dropping out [46]. During tutoring sessions in the physical learning environment, tutors obtain a substantial quantity of effective response which states the attitudes and feelings of the learner; whether she/he is happy, frustrated, confused, bored or surprised. The tutors use students' feedback to adapt their teaching strategies [133]. This suggests that this interaction should be replicated in distance learning environments; that they should respond to students' emotions, and accordingly guide and motivate them for better learning [133].

Many studies have pointed out that a learning system with sentiment interaction may raise study motivation and inspire learners during the interaction [166, 209]. Wang [209] designed a humorous and empathic virtual human to improve the existing e-learning system. The system reacts in a proper way, and fits students' emotions to assist students in overcoming negative emotions, and in enhancing students' learning motivation, interest, and performance.

Yu [223] addressed the different social skills of virtual companions in a VLE. The study aimed to simulate users' curiosity by constructing an intelligent Curious Companion (CC) to engage the users when studying science subjects. One of the aims of the current study will be therefore to identify from the psychology literature the most useful set of emotions to be used in a learning setting, together with the mechanisms by which they can be identified during a common interaction in a VLE.

Ferguson [59] highlighted the importance of understanding the capabilities of students through many tools such as activity-Based Assessment (ABA), where teachers keep monitoring students and assess them through a series of questionnaires. This profile helps teachers to develop the most effective teaching program that is based on students' interests. Similarly, the procedure followed by Ganotice [66] is to create students' emotion profile, then group students according to their emotion profiles. This profile reflected positively on the learning outcomes. The approach of using a questionnaire works with face-to-face learning, but less so with online learning, and it is cumbersome and costly to collect from time to time.

The availability of emotion recognition tools can automate this process and remove the burden of following up with a questionnaire. For instance Synesketch [104] was designed to track online communication and recognise emotions in the text; it was successfully integrated with Skype. The availability of tools that can be integrated with online learning systems and automate the process of emotion recognition can simplify the process of extracting and tracking students' emotions. Details of the detected emotions can be stored in a database to be analysed anytime using statistical analysis packages.

The analysis for this emotion can empirically provide the previously mentioned efforts of identifying the impact of students' emotions on their achievement, engagement and outcomes; valuable information that serves several purposes such as teachers and lecturers who are concerned with dividing students into groups based on their emotional profile and achievement as [66] suggests. So, using the emotional profile auto-generated by emotional recognition tools can considerably facilitate this process.

In this thesis, we will therefore focus on learning environments in which the majority of interactions happen via textual dialogues. This will offer more scope for generalisation, as textual interaction is the baseline for all VLEs. Sentiment and emotion analysis will be used to recognise the emotion when students use words to express their emotion and give their opinion [26, 162].

Our algorithm will combine both sentiment analysis and emotion classification: accounting for polarity scoring (whether the actor's sentences express a positive or negative attitude) and emotion states (happy, sad etc.) which will enable us to characterise students by their sentiment towards a set of notions that exist in the test corpus that are believed to be of specific interest [26, 162].

2.3.4 Detecting Sentiment in E-learning Context

Sentiment analysis is one of the fastest-growing areas of research in recent years, around 99% of the papers on this topic have been published since 2004 [123]. The crucial primary benefit of this automated process is to classify a given document or text based on its polarity - either positive, negative or neutral. It is also known as opinion mining; deriving the opinion or attitude of a speaker [113].

It could be said that Sentiments are feelings, attitudes, emotions, opinions, or subjective impressions, but not facts [113]. Sentiment analysis usually employs natural language processing (NLP), statistics, or machine learning methods to extract, identify, or otherwise characterise the sentiment content of a text unit. In general, sentiment analysis could focus on words, sentences, or documents. The earliest work on sentiment analysis, in the late 1990s, was performed on the classification of words or phrases based on semantic matters [77].

Sentiment analysis is typically used in marketing or social media trends, the early work in this area includes research started by the studies of Turney and Pang [149, 203] to detect the polarity of product reviews and movie reviews using machine learning techniques. A deep syntactic-semantic approach was used by Maks and Mohammad [122, 132] to develop a lexicon model for subjectivity description. This model tests the attitudes based on the verbs used in the sentences. Similarly, studies by [89, 95] determine the positive and negative terms of customer movie reviews taking into account a dictionary of synonym differences. Deep language analysis techniques employed in Kim and Hovy [97] still concern the availability of keywords and study their impact on the other words in the sentence.

One promising area of investigation is the use of sentiment analysis and its application to education. The study in [9] proposed a system to automatically evaluate sentiment from real time students' feedback in the classroom. The study applied data mining techniques to construct sentiment ratings, and found that lecturers were extremely satisfied with the system. Students sometimes used social networks to post their experiences. A recent study in [147] looked at extracting sentiment from messages the students write in Facebook. The study used the SentBuk application to retrieves the messages, comments and likes on the user's profiles and applied combinations of lexical-based and machine-learning techniques to extract the polarity.

Other researchers aim to enhance the communication between teachers and learners by analysing students' learning diaries [142]. Learning diaries are written containers, which are usually documented over a period of time [134]. Analysing theses records provides the lecturers with insights into what students feel and think during lectures.

In the context of e-learning, as most of the interaction is textual, the procedure of determining whether a piece of students' writing is positive, negative or neutral will be a valuable and interesting contributions. It is very useful to get information about the feelings of the students to support personalised learning. Several researchers investigate the implementation of Sentiment Analysis in e-learning systems.

An early work presented in [189] was to identify and extract the degree of the negative sentences and how this would be useful in an e-learning system. Similarly, the experiment in [23] proposed an opinion mining framework that can be applied to various domains, including education, especially when applied to an e-learning system.

The approach proposed in [94] classifies the reviews about an e-learning system as positive and negative, to adapt and develop the approaches and procedures used in teaching. That is, the techniques and processes of learning and teaching can be improved in some respects when the opinions of users are recorded and classified according to sentiment polarity.

The instructors benefit from this tracking by being able to detect any changes in students' emotions in an online environment, as in a face-to-face class, teachers can easily identify emotions in student facial expressions. In this context, the finding from study in [38], which proposed a framework for e-learning called "SAFE", showed that the teacher altered his teaching style for some topics in the course according to the mood of the class.

2.3.5 Detecting Emotions in E-learning Context

For several years, researchers in education have underestimated the role of emotions in a student's life, focusing exclusively on cognitive, motivational, and behavioural constructs [66]. Recently, there has been a growing recognition that emotions have a crucial impact on students' learning outcomes [154]. A growing body of literature has recently demonstrated the significance of emotions in e-learning processes by highlighting their impact on academic achievement.

It has been proved that emotions are common in online learning settings [14]. Perkun

[153] reported that emotions are common in the e-learning environment as students are shown to experience specific negative symptoms such as anxiety, boredom, exhaustion, and frustration during particular challenging tasks. There are a number of examples of studies that classified emotion expression automatically in texts [10, 46, 169, 194]. Thus, several authors claim that emotions could be detected from different online learning contexts (e.g. forum discussions, chat discussions) and that can be formal with a teacher or informal amongst students peers [53]. A study by [46] emotions are associated with students' engagement. The e-teacher should aim to reduce the negative emotions of the students, especially the negative emotions identified during the chat interactions with teacher or students' peers.

The "control-value theory of emotions" [155] links emotions and achievement motives, activities, and outcomes. The theory argues that emotions determine key learning processes [66]. Students who experience positive emotions such as enjoyment and pride are expected to achieve higher grades [156, 208], engage in the more effective use of cognitive and meta-cognitive strategies [15, 98, 141], and participate more actively in-class activities [66, 98, 99, 156]. Similarly, as expected, the negative emotions (such as boredom) are found to lead to low achievement [157] and decreased effort [47]. A study by [46] found that negative emotions play a central role, as detecting negative emotions across the e-learning activities may be the first indication of insufficient learning.

Various studies investigated emotions for many purposes, such as seeking customers' perspective for new products or services [181]. Some used text-mining techniques [222], others proposed a model for statistical analysis of collective emotions in the text [67]. The predominant approach to emotion classification is based on the premise that the overall emotion of a sentence is the aggregation of the sentiment of the words comprising it, as well as syntactic and semantic information [181]. These techniques seek the presence of appropriate affect words in text. Some words are unambiguously affect words [181], while others carry affect to some degree. This method either uses a corpus-driven approach to give affective orientation or scores to words, or it depends on some existing affect lexicons.

Various emotions are experienced by students during the study depending on the situation. For instance, the student may become happy when they score a high grade, and upset when they fail. Many studies investigated various types of emotions. Pekrun [156] developed an Achievement Emotions Questionnaire (AEQ) to assess multiple achievement emotions experienced by students in academic settings. The emotions investigated are enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, and boredom. Those emotions were investigated during class, while studying, when taking tests and exams through the questionnaire.

Krithika [106] used "excite", "disturb" and "moving" pattern of eyes and head to deduce meaningful information to understand the mood of the student when engaged in an e-learning environment. However, they did not look for specific emotions but only aimed to recognise that the mood is either positive or negative. With the same context Litman and Forbes-Riley [112] predicted emotions from the voice. However, they only categorised emotions into positive, negative or neutral. The six emotions recognised within the adopted tool used in our experiments (Synesketch tool) [1] are anger, disgust, sad, surprise, happy, and fear. Those six have been recognised by [156] and many researchers followed this approach with minimal changes. In our third experiment [6], more emotion types are recognised using the Emotion Clock which was introduced by Whissell [214] to categorise emotions based on Russell's model dimensions [171].

Emotions in text can be detected using various approaches and methods, Naresh reported [8, 182] that there are four methods of text-based emotion detection, see Figure [2.6]. They are Keyword Spotting Technique, Lexical Affinity Method, Learning-based Methods, and

Hybrid Methods. Keyword Spotting can be defined as the searching techniques to find the occurrences of keywords in utterances.

In term of emotions detection, keyword spotting deals with finding the relevant instances of words based on the predefined set of keywords. The keyword spotting mechanism is shown in Figure [2.7] where text is considered as input, then converted into tokens, each word related to emotions is detected and identified, and finally an emotions class obtained as output. Our adopted tool in the first experiment (***Chapter [4]***) in this thesis, Synesketch, uses the Keyword Spotting algorithm to detect emotions in text.

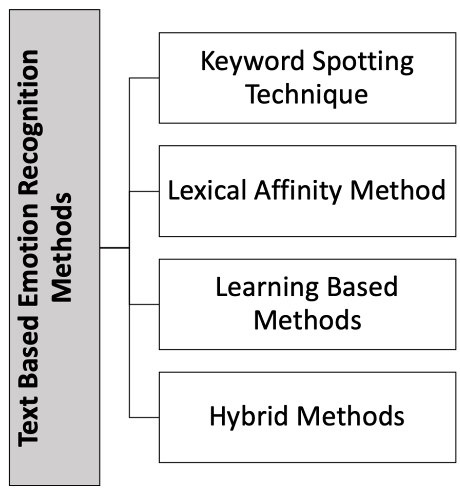


Figure 2.6: Several methods of emotion recognition based on text [182]

Lexical Affinity methods are a straightforward and easy to use process where emotions are identified based on related keywords and a constructed emotion lexicon. Our second adopted algorithm in the third experiment (***Chapter [6]***) to identify emotions is based on a lexicon where we use the dictionary to ensure we have all the synonyms of an emotion word.

The third method to detect emotion is a Learning Based Method where the detection of emotions is based on various theories of machine learning classifiers, such as support vector

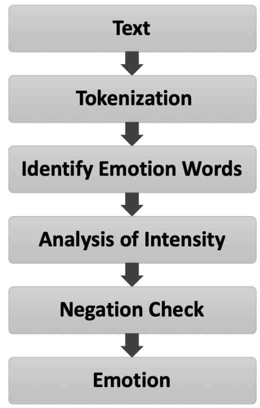


Figure 2.7: Keyword Spotting Technique process [182]

machines. Finally, the hybrid Methods which is based on joining the two algorithms: the keyword based method and the learning based method.

2.4 Stylometric (Writing-style) Features

Human beings have the cognitive ability to understand different styles of writing and to distinguish between different authors' writing samples. With the advent of computers and the growing body of texts available via the internet, *stylometric* analysis is used as a technique for analysing texts [138]. The development of the stylometry in early 1890 was lead by the Polish philosopher Wincenty Lutoslawski. Writing-style features are considered as the primary characteristics of stylometric analysis [228].

Many applications of stylometric are for identification and similarity detection; it can be also be used in the early detection of diseases [1, 45, 80, 120]. Many researchers, especially in natural language processing, consider stylometric features for evidence of authenticity to help identify individuals based on their writing style [2].

There are different types of writing styles features [1] lexical (e.g. character level, word level, and richness) [225, 228], syntactic (e.g. punctuation, total number of words, and Part of Speech - POS) [138], structural (e.g. how many sentences, url, greetings, email structure) [45], and content-specific level (important keywords and phrases on certain topics) [125].

Understanding the importance of syntactic and lexical information remains an open research area in linguistics [24, 109]. Previous works on writing style analysis hint that features could be used for various reasons. A study in [196] proved that writing style features extracted from posts in social media could be used to objectively characterise user reputations. They analysed writing styles to estimate if the author changed their writing style or not. They trained eight classification techniques in order to detect the changes. The authors characterise "Good" when their writing does not change (no mistakes, perfect paragraph, structure good or not, ... etc.), and "Bad" when changes occur.

In another study in [2], linguistic features were used to investigate the forensics of the user. They used ML techniques to detect the identity of the author in any new post by them. They needed to see if an author changed his/her writing style in comparison to their last post. In addition, they can check to see if an author copies the style writing of other authors. A study in [69] measured the changes over time in the writing style of seven authors of novels written in English. The work included three types of stylometric analysis: phraseology (e.g. mean word length, mean sentence length), punctuation (e.g. commas, semicolons, quotations) and lexical (e.g. stop word list) usage, in order to compare three stages of writing for each author.

Each of these writing styles is used for a specific purpose: vocabulary richness is a natural way to detect if the person uses different words than those that they would normally use in the writing [224]. Using different words indicates how well a learner can convey meanings and express their writing. It plays a central part in judging the quality of a student's writing, because a learner corpus may contain a very high rate of non-standard forms of writing [72]. Richness is also used in [221] to identify if the person is the actual writer of the content posted in social media posts or if it is someone else.

One of the most important features is to look at syntactic information. Part of speech (POS) is an essential aspect of syntactic information, as it assumed by many researchers that an author of a text could subconsciously use similar syntactic structures to write sentences [17]. The finding in [2, 196] proved that individuals who are using correct grammatical tagging are more confident in their writing.

In this thesis, we determine how the writing styles of students can be used as an objective feature for estimating students behavioural change over-time. We are trying to understand the effect of different features such as Lexical and syntactic features: the vocabulary richness and Part of Speech (POS) of the text.

2.5 Conclusions

A review of the literature addressed some of the concepts related to the research presented in this thesis. The chapter presented recent approaches and contributions to the role of sentiment/ emotion analysis in students' engagement in the e-learning environment. In all types of communication such as speech, text, and gestures, emotions and sentimental states are inescapable. The emotion/sentiment definitions, origin, algorithms and techniques have been examined. There are various strategies, software solutions and intelligent approaches developed for educational purposes. The literature showed the role and the use of emotion analysis in assisting to find approaches that address the problem of disengaging students, or students struggling with the online learning, and how educators can use this information to adjust their teaching.

Also discussed in the literature review were issues associated with the challenging task of tracking students in the online environment. Learning Analytics are the key tool to monitor the learning process and students' progress, and to try and prevent student disengagement issues, however, they present issues, which have been discussed, most notably that LA tools concentrate on statistics of engagement, and often ignore the context in which these statistics come about, for instance the topic of conversations, the analysis of which is still a burden for e-educators.

We also gave an overview of techniques to analyse textual messages, and this gave us the opportunity to discuss the importance of the use of writing style, by discussing the concept and the primary characteristics of stylometric analysis, which are important to investigate the students behavioural changes over-time.

This survey is by no means exhaustive, and more pointers will be given in the following chapters, when discussing specific tools and techniques, as they are useful to explain more technical aspects of the experiments presented in this thesis.

In the next Chapter, we will present the conceptual framework produced as a result of our investigations, and will describe the approach and strategies used for the evaluation of this framework.

Chapter 3

Conceptual Framework and Methodology

3.1 Introduction

As mentioned in the previous Chapters, we are witnessing a mental health crisis in Higher Education, with researchers reporting that students are increasingly at risk of dropping out because of lack of mental health support, more so than because of academic issues.

Ideally, students at-risk should be identified as early as possible to allow sufficient time for instructors to conduct educational interventions to facilitate students' learning achievements. Identifying at-risk students in the context of online learning is a challenging and demanding task, and there are a variety of strategies and systems to facilitate students in the learning process and minimise the impact of negative behaviours.

The development of a sophisticated "system" able to monitor human emotions and react properly to the user's behaviour is of course a promising area of research. However, such technology is often complex and relying on devices and sensors which detect the physical manifestations of human emotions. This is not always, in fact it is very rarely, a realistic scenario for online-learning, where students, and sometimes also teachers, participate to the learning experiences from their own locations and with their own devices.

Moreover, in deciding what type of digital intervention to design, it is important to establish the role of such system, especially with respect to how it is positioned in the interaction between teachers and learners. Our main objective is to improve the dialogue between teacher and student, rather than substituting to the teacher, and therefore we envisage a system that could be integrated in the common interaction of a virtual classroom, and hence help both student's learning and lecturer's understanding of and dealing with a student's motivational issue.

This chapter aims to establish the conceptual framework that would support the design and implementation of such a system, by discussing the various components and the role each of them plays towards the final goal of being able to help educators better identify the issues that need intervention. The Chapter also describes our strategy for evaluating this framework, in a series of feasibility studies which are then presented in detail in the following Chapters.

3.2 Conceptual Framework

The main contribution of this thesis is the formulation and evaluation of a conceptual framework for supporting the implementation of a system able to identify students at risk, in such a way that the role of the online tutor or facilitator is incorporated into an environment that coordinates information coming from the various communicative acts happening during learning, from tasks in the classrooms, to question to the teachers to discussion among peers. A pictorial representation of the conceptual framework is in the Figure below [3.1]. The framework consists of the following components:

Educational Platform. We assume the presence of an educational digital platform, as a collaborative interactive online learning environment where teachers, learners and other stakeholders use tools and resources to support and improve the delivery of education. The learning platform contains list of features and functions (tools) that facilitate and support the learning process, such as lecture notes, discussions boards, quizzes, assignments, audio/video lessons, virtual classrooms and others services. All the information and resources are shared online over the Internet. We assume that the users of the platform may have different roles, from what can be identified to a "lecturer", or academic staff responsible of a specific unit of learning, to pastoral advisors and mentors, to academic or administrative coordinators of a series of units, or a degree.

Prediction System. Alongside the educational platform, we assume the presence of a "Prediction System". The prediction system would be based on the different methods and techniques available to track students' needs and progress, and will be equipped with algorithms for the analysis of textual messages, as well as any variables associated with students, from their performances and general learning analytics, to their personal profile. The prediction system should be able both to follow the journey of a specific student, and to compare this journey with other similar journeys of the same cohort, or previous cohorts, in order to offer a better picture of how some behaviours can be thought as "the norm" for a given module or a given situation. We especially envision a prediction system that is not based on pure engagement, as there are many in the literature, but which pays special attention to the emotional state.

The purpose of this system, rather than acting on behalf of a lecturer, would be to raise awareness on students' performance, so that a human stakeholder can intervene at the appropriate time. The communicative acts are all "observed" through the lens of the "emotional loaded" prediction system, which can therefore decide to intervene, or simply flag, when some communication needs to be looked at more carefully.

The conceptual framework proposed inn this thesis is evaluated by a series of experiments and case studies, aimed at assessing feasibility and appreciate usefulness of single potential components, and on the basis of expert commentaries, as described in the following section.

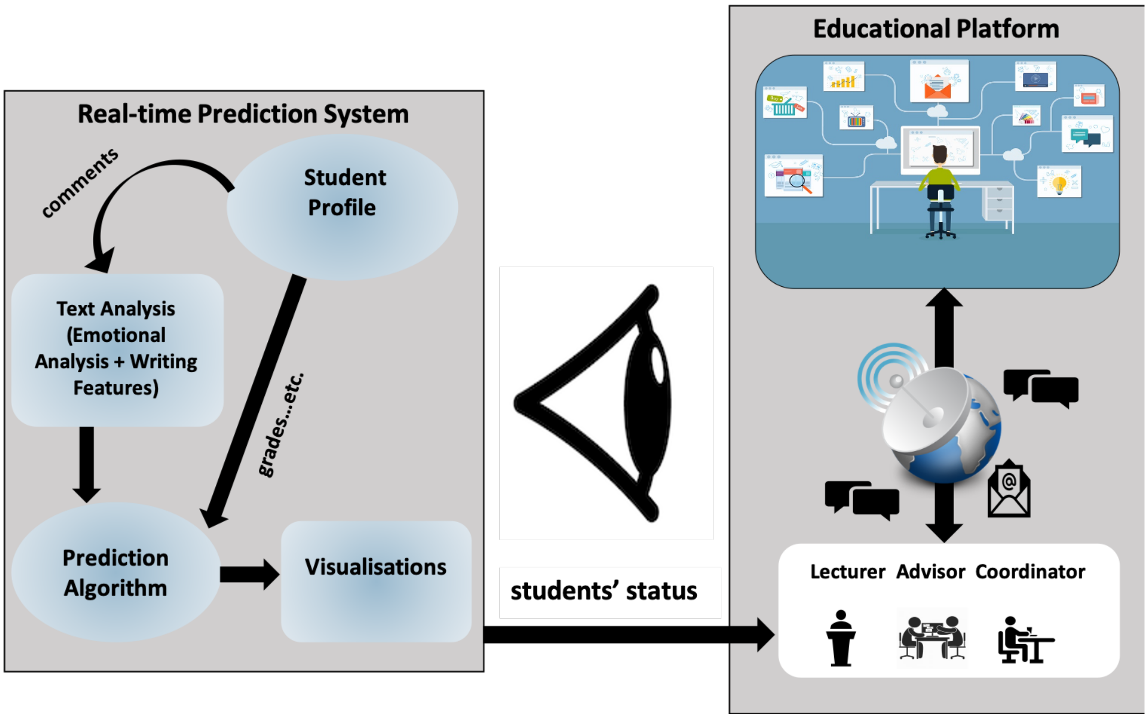


Figure 3.1: Structure of the "observer" system

3.3 Feasibility Study: Methodology and Experiments

One of the main challenges in educational research is evaluation: ethical guidelines offered by associations, such as the British Educational Research Association we adhere to [28], insist, and rightly so, on the notion of consent, such that it would be normally expected that all participants, teachers and students alike, involved in a study ought to be able to confirm their willingness to take part, and at the same time to withdraw their consent at any time and for any reason. While this is not a big obstacle for smaller interventions, the evaluation of a complete framework such as the one we propose in this thesis poses a much bigger circular challenge: we are clearly not in a position to deploy such a complex system in a real learning environment for testing before we can evaluate it, as this would involve a large commitment from an active learning provider, but we cannot evaluate it in a realistic environment without full deployment.

In this thesis we provide a comprehensive and realistic full feasibility study for a system such as the one described in Fig. 3.1. We approached the task by breaking it down into a series of feasibility studies where we perform an evaluation of single components or tasks that we would like such a system to perform. The main approach to this is based on the following principles.

3.3.1 General Principles

1. **Practicality.** When evaluating a task, **we rely on off-the-shelf systems** or algorithms which are open source, and we use them in such a way to answer the question: *is this task achievable*? We do not therefore seek to carry out a performance study, in order to establish which tool would be best efficient for the task, or which will deliver the optimal performance. The reason for this is that we are focussing on feasibility only, in the assumption that, should one wanted to deploy such a system in a real environment, a performance comparison would be on order, with performance of the various algorithms measured for the specific of the learning platform. In this way we can remain platform-agnostic and concentrate only on said task feasibility, while we acknowledge that the solution demonstrated is not necessarily an optimal one.

2. **Reproducibility.** When performing an experiment, **we rely on readily available,**

general datasets, rather than creating one *ad-hoc*. For each of the task proposed we identify a dataset which can demonstrate both the value and the limit of the solution, and can offer an insight on whether current educational datasets are rich enough to support systems that go beyond the simple learning analytics. For example, on one experiment in particular, we found that no educational dataset was suitable, and we had to resort to an alternative solution, as we show later in this document. This was in itself an interesting finding, and serves as recommendation for providers who may want to consider adopting a framework such as the one we propose.

3. **Practitioner validation.** To maintain an objective view on our analysis, **we rely on expert opinion** which we captured through a focus group study at the end of our research. By expert we intend practitioners in the field, online educators and advisors, who are the most likely to use a tool of this kind, should it become available, and can comment on the tool uptake and critical issues to consider.

In the remainder of this section, we summarise the set of experiments which will form the content of the following Chapters of this document.

3.3.2 Map of Experiments

We devised and carried out a set of four studies in order to evaluate the conceptual framework, consisting of three experiments using computational tools applied to relevant datasets, and one focus group study, to collect feedback on the proposal.

Experiment 1 looks at the left hand side of the conceptual framework picture, and tools for the analysis of emotions in textual messages, in order to understand how feasible it would be to put together a tool able to detect emotion in students messages. We apply this to a dataset of self reported emotional episodes accounts, which simulates potential personal messages for instance to their lecturers or advisors (***Chapter [4]***).

Experiment 2 looks at part of the left hand side of the picture, and employs tools which are able to use emotion/sentiment analysis as a predictor of students' final grades. We apply this to a dataset coming from an online education provider, which also gives us an understanding of what type of student profile is commonly provided by such datasets (***Chapter [5]***).

Experiment 3 goes a step further in the analysis of the student profile by exploring the writing style of a student, as well as the emotional content of messages, and furthermore consider how we can look at a whole classroom from the point of view of the observer, and flag cases for potential follow up by the lecturer. The idea is to form a baseline for each student that can help identify anomalies *with respect to this student's norm* as opposed to generic emotion analysis (***Chapter [6]***).

Finally, with **Experiment 4** we describe the outcome of a focus group study, that we carried out by inviting long term experts in e-learning and online class facilitation, to gather feedback on our approach (***Chapter [7]***).

3.3.3 Datasets for the Experiments

One of the principles we used for our study was reproducibility, which meant we insisted on relying on existing general datasets for our experiments. This meant that we were not able to put together our ideal data, which would have included a set of observations from a number of real courses, delivered to the same cohort of students, and showing their progress from one module to the other, as well as discussing in open text any issues related to their degree and studies. Collecting such a dataset would pose a number of ethical challenges, as well as being not feasible in the time frame of this thesis. We had therefore to compromise and settle on datasets to use which were *suitable enough*. This is explained in detail in each chapter, but we provide a summary here.

The dataset the first experiment (***Chapter [4]***) had been collected from the International Survey On Emotion Antecedents And Reactions (ISEAR) [176]. The main advantage of this dataset is that emotions were self-reported by the participants, therefore we were provided with a baseline to check our chosen emotion analysis system against. Also, participants to the creation of the dataset were students, and topics were related to an educational situation. The main disadvantage was sentences in the dataset were generic and not related to a specific class, or class situation, so they would not give us the sense of a cohort of students together in the same journey.

In the second experiment (***Chapter [5]***), we concentrate on actual classroom behaviour, and we have opted the dataset from the Stanford MOOCs dataset, as made available by the Center for Advanced Research through Online Learning (CAROL) [35]. The main advantage is of course the setting: this is real data of students in a classroom. The disadvantage is that the MOOC setting did not encourage open text discussion, therefore we are limited in the amount of text we can analyse.

We used for the third experiment (***Chapter [6]***) a Motivational Interviewing corpus of transcripts [5], which consists of a searchable collection containing real transcripts of counselling and therapy sessions between client and counsellor. This is probably a more controversial choice, as the corpus is neither coming from an educational setting nor is coming from a group of participants on the same session. However, the corpus is extremely reach in emotional loaded text, and does provide the notion of a journey, with the patient and the counsellor working towards the same goal. The corpus was adapted in a series of ways, explained in the Chapter, which made it possible to create an analogy between:

- the patients and the students,
- the counsellor and the teacher/facilitator of a classroom,

- the set of sessions for the same patient and the "lectures" of a classroom
- and the set of patients discussing with the counsellor on a given issue and "a cohort/class" of students working with the teacher on a given topic.

With this analogy in mind, we are able to talk about "emotion of the cohort" and this helps us identifying possible strategies to calculate and to visualise this notion to a teacher.

3.4 Conclusions

This chapter presented the conceptual framework proposed by this thesis, by also providing the pedagogical motivation behind the proposal, and outlined the approach that we used for a systematic study which addresses the issue of feasibility of the approach. The following Chapters will detail the experiments identified above. For each of these experiments, the relevant Chapter will describe motivation and methodology, then will describe the tool that has been used to carry out the experiment, then the dataset that was used, and finally the outcome of the experiment. A final discussion on the lessons learnt and the limitation of the experiments will conclude each Chapter.

Chapter 4

Experiment 1: Capturing Emotions in Student's Narrations

4.1 Introduction: Experiment 1 Overview

One of the aims of our research is to explore the use of emotion analysis to help enhance the relationship between learner and teacher in an online learning environment. The idea, as described in the previous Chapters, is to complement the role of the online educator, by providing information about behavioural patterns and emotional states of the student, in order to eventually suggest strategies for intervention.

Therefore, one of the first tasks that we wanted to evaluate is emotion identification. More specifically, the aim of the experiment was to establish the extent to which a generic off-the-shelf emotion analysis tool can be used in the context of messages exchanged in an educational context, not only messages related to the classroom, but those that, in our conceptual framework, arise from the interaction with other actors, like mentors or advisors, or administrative staff.

In the following sections, we will describe the tool that we used for the experiment, motivating our choice, and the dataset we selected for testing. We then show the results of the experiments, and we discuss the significance of these.

4.2 Emotion Capturing Tool: Synesketch

For the first experiment, we utilised a tool called Synesketch, an Open Source Library for Sentence-Based Emotion [104], which has already been mentioned in the literature review in Chapter 2. The main advantages of this tool is that it is free, open source, and can be integrated with exchange messages systems like Skype. The tool has been successfully employed in a number of projects related to emotion extractions, such as [4, 70, 174]. Also, it is simple to use by non experts, and therefore, while it is by no means the most sophisticated emotion analysis tool, nevertheless it offers a useful baseline for testing what is achievable and possible.

Synesketch is based on lexicon, and uses a hybrid method of a keyword-spotting and, a rule-based method [174]. Keyword spotting is an approach for searching relevant known keywords in databases [91], for which many algorithms have been proposed in the past, and several techniques claim to be very efficient and accurate. A Keyword-spotting approach here is based on the use of a lexicon of words and expressions related to emotions. A word lexicon is utilised to search WordNet®[58, 131][1] for all those terms in the initial word set that are semantically related to the words in the lexicon (semantic relatives). Common abbreviations and emoticons such as ":)"s, ">:O"s, and "ROFL"s etc. are included. There are six emotional weights attached to each lexeme (word or emoticon), that relate to the six basic emotional categories identified by Ekman [54]: *Happiness, Sadness, Anger, Fear, Disgust,* and *Surprise.*

[1] `https://wordnet.princeton.edu/download`

4.2.1 Synesketch Tool: Algorithm Steps

The approach on Synesketch is tailored to the specific issue of analysing small fragments of text like online comments, tweets, short messages, etc. and is based on keyword spotting combined with affinity lexicon, and some heuristic rules [104].

The keyword spotting technique is based on a lexicon consisting of two parts: a word lexicon, generated semi-automatically using WordNet, and an emoticon lexicon, using common abbreviations and shorthands for emoticons (e.g. ":-)") [104]. Each lexicon entry is associated with one of the six emotional categories, with a value between [0,..,1]. More details on how this was built are in [104].

The emotion classifier algorithm consists of the steps in Algorithm 1 (see also Fig. 4.1). When applied for example to the sentence "I won't be lovesick!" [104], the algorithm will spot the keyword "lovesick" which has the emotional vector: $[0, 0.9, 0, 0, 0, 0]$, therefore has only a value for sadness at 0.9. The sentence level rules will recognise that there is a negation "won't", so the negation change the value of sadness to happiness. The happiness weight takes the value of the dominant negative weight, which is only one in this case (sadness), so the new vector becomes $[0.9, 0, 0, 0, 0, 0]$. Finally, because the sentence ends with a "!", the intensity is increased by 20%, and the vector becomes $[1, 0, 0, 0, 0, 0]$. The emotional valence is 1.

4.3 Dataset for Experiment 1: ISEAR

For this experiments, we wanted to establish how the selected tool could perform in an educational setting, by identifying emotions which are likely to be aroused in students in specific educational situations. In order to do so, we needed a dataset which had been especially collected among students. The natural place where to search for such datasets

Algorithm 1 Synesketch general algorithm (from [104])

1. The input sentence is interpreted by applying sentence-level rules, related to negation and punctuation.
2. The input sentence is parsed into words; each word is matched with both lexicon keywords.
3. When a keyword is spotted, the emotional weights of the keyword are considered, and adjusted on the basis of other sentence level rules (e.g. if the emotional keyword is uppercase, then the value is intensified by 50% etc.).
4. The keyword is added into an emotion word set. This step is applied to all spotted emotion-related keywords.
5. The sentence's total emotional state, that is the overall vector corresponding to the whole sentence, is calculated using the emotion word set with updated weights.
6. The overall emotional weights of the vector depend on the maximum value among all the weights of the keywords of the same emotion type from the emotion word set.
   (a) If the general weight is 1, it means the Synesketch is highly confident in determining one or more emotions.
   (b) If the general weight is 0, it means Synesketch failed to identify an emotion. In this case, we consider the sentence emotion neutral.
7. An emotional valence is also calculated, depending on whether the weight aggregation of the overall happiness outweighs the overall weight of the dominant negative emotion (sadness, anger, fear, or disgust).

is among those offered by online education providers. However, such datasets are mainly related to classroom interactions, and therefore do not offer the rich variety of situations which we wanted to simulate, for instance when students contact their advisor or mentor to express a particular concern.

We found a suitable dataset for our purposes in the International Survey On Emotion Antecedents And Reactions (ISEAR), which is a dataset collected in a project led by Klaus

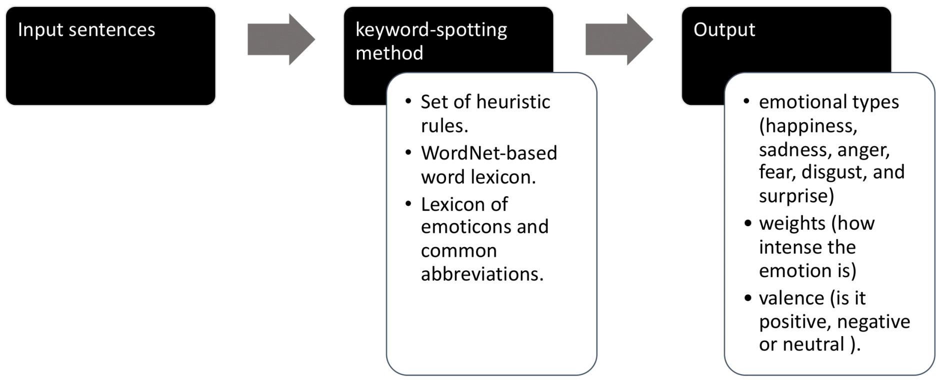


Figure 4.1: Structure of the emotional capturing tool (Synesketch tool)

R. Scherer and Harald Wallbott [176][2].

The dataset was put together over a number of years with student participants, who were asked to recall situations in which they experienced all or some of seven major emotions (*Joy, Fear, Anger, Sadness, Disgust, Shame,* and *Guilt*), by also reporting how they appraised the situation and how they reacted. The dataset comprises these recollections from around 3000 respondents, in 37 Countries from all 5 Continents. We chose this dataset as it was likely to include, given that the respondents were students, sentences related to an educational situation. As the experiences were self-reported, and not captured during learning activities, the dataset could give a reasonable simulation of the range of emotions students are likely to report to experience alongside their learning.

[2]Available for download at `https://www.unige.ch/cisa/index.php/download_file/view/395/296/`

4.3.1 ISEAR Dataset Preparations

For our study, we explored the dataset, and manually extracted around 1050 records related to education scenarios, such as records that described students' emotions with respect to their educational performance, their fellow classmates, their lecturers, the institution they study at, and exams, while of course reporting of various circumstances around the episode. Overall these give a good sense of the type of matters students would want to discuss, for instance, with their academic support. A small sample of those records are shown in Table [4.1], while for a more comprehensive set refer to [A.1] in Appendix A.

Text from ISEAR dataset	**labelled Emotion**
I was selected to come here (University, College) when I was least expecting it.	Joy
A neighbour's girl had disappeared and many people were looking for her. Someone had gone to notify the police. Something had certainly happened to her.	Fear
A classmate told me I must have bribed the class leader to let me go to your English lecture.	Anger
When my car froze, and I could not start it.	Sadness
When I was about to clean the draining board and saw it looked a underneath the sink (I live in a students hostel).	Disgust
Once I was not ready for a seminar and I was asked to leave.	Shame
When I finished a love affair where I was responsible of the sad end.	Guilt

Table 4.1: Sample of the records from ISEAR dataset

	Anger	Disgust	Fear	Guilt	Happiness	Joy	Sadness	Shame	Surprise
Synesketch	✓	✓	✓			✓	✓		✓
ISEAR	✓	✓	✓	✓	✓		✓	✓	

Table 4.2: Emotions in Synesketch and ISEAR

4.4 Experiment 1: Using Synesketch on the ISEAR Dataset

In our experiment, we ran the selected set of records from the ISEAR Dataset through Synesketch, and we compared the results from the tool, in terms of the emotions detected, with the emotion type that was labelled in the ISEAR dataset, as self reported by the participants. It has to be noted that the two sets of emotions do not match entirely: while both ISEAR and Synesketch consider *Fear, Anger, Sadness and Disgust*, the main positive emotion is named *Happiness* in Synesketch and *Joy* in ISEAR (for our experiment we assumed we could equate the two), and more importantly, Synesketch includes *Surprise* as a sixth emotion, while ISEAR comprises *Shame* and *Guilt* as two extra, separate, negative emotions (see Table 4.2 for a summary). This provides us with another level to the evaluation, as we can presume it would not be uncommon for an off-the-shelf tool not to match entirely the specification of the dataset.

In this section we will discuss the results of this experiment, both from a quantitative and from a qualitative point of view.

4.4.1 Quantitative Analysis

When analysing a sentence, in addition to the emotion type, Synesketch also provides the weight, or intensity of the emotion, and the valence (whether it is a positive, negative or neutral emotion). An example of the output is in Table [4.3] and for a more extensive results refer to Table [A.2] in Appendix A. In the table, we report: the text form ISEAR, the labelled emotion from the ISEAR dataset, so the emotion the participant self reported, then a set of Synesketch results: the General weight of the emotion, representing how confident the tool is in determining the emotion (where 1 is the maximum confidence, and 0 if the tool failed to identify any emotion), the Valence of the general sentiment, where -1 is Negative, 0 is Neutral and 1 is Positive, and the 6 values the tool returns for each of the emotions.

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
When I passed the TOEFEL with very good marks.	Joy	1	1	1	0.2	0	0	0	0
When abroad, while driving a car along a dark, winding road.	Fear	1	-1	0	1	0	1	0.1	0
A classmate told me I must have bribed the class leader to let me go to your English lecture.	Anger	0	-1	0	0.6	0.4	0.6	0	0
The death of a close friend.	Sadness	1	-1	0	1	0	0	0	0
When I was told that a good friend was seriously ill.	Sadness	1	-1	0.2	1	0	0	0	0.3

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
My room-mate was drunk, he vomited on the floor and fell face.	Disgust	0.2	-1	0	0	0.1	0	0.5	0.1
When I was about to clean the draining board and saw it looked underneath the sink (I live in a students hostel).	Disgust	0	0	0	0	0.1	0	0	0
I experienced long ago when I was sightseeing Bulgarians in a foreign language	Shame	0.1	-1	0	0.1	0	0.1	0	0

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
When I finished a love affair where I was responsible of the sad end.	Guilt	1	-1	1	1	1	0.2	0.2	0

Table 4.3: Sample of the results by Synesketch

The number of sentences across all records extracted from ISEAR which we ran through Synesketch amounted to 1598 in total. The emotion label that was assigned by the tool was compared to the label assigned in the database by the participants self reporting their emotion. The overall accuracy, for the five emotions that were common to both sets, is 83.72%. The breakdown for all the five emotions is presented in Table 4.4, where the accuracy of Synesketch tool degrades when dealing with disgust as the tool recognised some cases as fear or sadness. For example, "When I was about to clean the draining board and saw it looked underneath the sink (I live in a students hostel).", the tool gives the higher polarity for sadness as it could not catch any keyword that refer to disgust.

Emotion	**Sentences Correctly Labelled**	**Accuracy**
Anger	388/448	86.60%
Sadness	332/389	85.34%
Fear	245/286	85.66%
Happiness	197/226	87.16%
Disgust	176/249	70.68%
Overall	1338/1598	83.72%

Table 4.4: Accuracy of Synesketch labels of the ISEAR sentences

As to the emotions that are not in common between the two settings, we have analysed the two cases separately.

4.4.1.1 Synesketch Surprise

Synesketch is programmed to identify "Surprise" which however is not present in the ISEAR dataset as a self reported emotion. The tool labelled 74 sentences with "Surprise" and Table 4.5 shows the distribution of categories they were labelled with:

Emotion in ISEAR	No. of cases labelled "Surprise"
Anger	14
Sadness	4
Disgust	20
Fear	24
Joy	12
Total	74

Table 4.5: Distribution of the ISEAR sentences that Synesketch labelled as Surprise

4.4.1.2 Shame and Guilt

The dual situation is the one regarding emotions *Shame* and *Guilt*: these are present in ISEAR but Synesketch is not equipped to identify them. These will obviously be classified differently, but it is interesting to know whether the tool still manages to identify them as negative emotions, and which emotion in particular. A summary of results is in Table 4.6.

Emotion in ISEAR	Classification by Synesketch	Accuracy
Shame	62/273: Anger	
	44/273: Sadness	
	38/273: Disgust	
	0/273: Fear	
	Total Negative: 144/273	52.74%
	Total Neutral: 88/273	32.23%
	Total Positive: 41/273	15.01%
Guilt	47/290: Anger	
	73/290: Sadness	
	10/290: Disgust	
	24/290: Fear	
	Total Negative: 154/290	53.10%
	Total Neutral: 94/290	32.41%
	Total Positive: 42/290	14.48%

Table 4.6: Synesketch labelling of ISEAR Shame and Guilt

4.4.2 Qualitative Analysis

Synesketch showed high accuracy when the sentences are short and contain keywords with an emotional load. However, with long sentences, the tool is not as accurate, as one might expect with off-the-shelf solution. In this section we will discuss some of the problematic cases, and what they can teach us on a way to implement a tool able to capture the more complex students' sentences. In what follows we will use the notation **"Sentence$^{\textbf{(Emotion)}}$"** to indicate a Sentence from ISEAR that was labelled by the participants with an Emotion.

The most common outcome for long sentences is that the tool is unable to decide which emotion is prevalent and classifies them as Neutral. A good example is sentence **"When I heard on the radio that the football match in Belgium had ended in a catastrophe$^{\textbf{(Anger)}}$"**, which was classed as Neutral despite the presence of the very loaded keyword 'catastrophe'. The problem is clearly related to the length of the sentence, as for instance sentence **"When I had been obviously unjustly treated and had no possibility of elucidating this$^{\textbf{(Anger)}}$"** and was labelled Neutral, even though the word "unjustly" is a word with negative association, however the sentence **"When one is unjustly accused of something one has not done$^{\textbf{(Anger)}}$"** was correctly labelled.

The tool seems to weigh nouns more than verbs, so, for instance it is more likely to classify accurately a sentence containing the noun "death" than one with the verb "die": sentences **"I heard that a former superior of mine had died$^{\textbf{(Sadness)}}$"** and **"When I came to know that my grandmother had died$^{\textbf{(Sadness)}}$"** where both considered Neutral. However, sentence **"My grandmother died suddenly last summer$^{\textbf{(Sadness)}}$"** the emotion "Surprise" was the highest prediction with 100% of confidence (possibly because of the presence of "suddenly") and "Sadness" was far behind with 30% of confidence, and sentence **"When one of my cat died of a disease$^{\textbf{(Sadness)}}$"**, was correctly classified, possibly because of the presence of "disease". The most notable example is for the sentence **"I realised that the girl I loved had to leave me$^{\textbf{(Sadness)}}$"**, which had a high score for happiness at 0.80%.

The emotions in ISEAR but not part for the tool, guilt and shame, were most often labelled with anger, fear, and sadness. For instance, for sentence **"I feel guilty when**

when I realize that I consider material things more important than caring for my relatives(Shame)", the tool picked fear and disgust as the top emotions. Similarly, with this sentence, **"When I realized that I was directing the feelings of discontent with myself at my partner and this way was trying to put the blame on him instead of sorting out my own feelings(Shame)"**, the tool predicted anger as the top emotion, followed by disgust.

The tool is particularly problematic with long sentences, often resulting in no prediction (e.g. for **"It was more the ceremonies held in the church and not really the death of my grand-father which made me sad(Sadness)"** the tool equally predicts negative and positive emotions), and sentences including a negation (e.g. **"not being able to marry and have children(Sadness)"** and **"I was told by a good friend that we couldn't be friends any more because of his relationship with another girl(Sadness)"** had both an overall positive feeling), as well as, as expected, figurative language (e.g. **"They called her a pig and then grunted(Anger)"** was considered void of emotional load).

Other problematic cases are presented in Table [4.7], and a more extensive set in Table [A.3] in Appendix A.

Sentence	Failure Categories	General Weight	Valence	Anger	Disgust	Fear	Sad	Happy	Surprise
When my friends did not ask me to go to a New Year's party with them.	problem with negation	0.1	1	0	0	0	0	0.1	0
When I read a theoretical book in English that I did not understand.	problem with negation	0	0	0	0	0	0	0	0
When one's studies seem hopelessly difficult and uninteresting.	keywords not recognized	0.6	0	0	0	0	0	0	0
When I was not accepted as a student in finance and accounting.	problem with negation	0	0	0	0	0	0	0	0

Sentence	**Failure Categories**	**General Weight**	**Valence**	**Anger**	**Disgust**	**Fear**	**Sad**	**Happy**	**Surprise**
When I had not understood anything after a lecture.	problem with negation	0	0	0	0	0	0	0	0
A case of unrequited love.	problem with negation	1	1	0.1	0	0	0	1	0
Thoughts revolve around failing the subject and the consequences it would have for the future.	keywords not recognized	0.1	1	0	0	0	0	0.1	0
Not succeeding in a cross-country skiing competition;	problem with negation	0	0	0	0	0	0	0	0

Sentence	**Failure Categories**	**General Weight**	**Valence**	**Anger**	**Disgust**	**Fear**	**Sad**	**Happy**	**Surprise**
When I did not get the salary increase that I had been expecting and understood how little one's work was appreciated.	problem with negation	0.1	1	0	0	0	0	0	0
I heard that a former superior of mine had died;	keywords not recognized	0.4	1	0	0	0	0	0.1	0

Table 4.7: Sample of failure cases of results by Synesketch tool

4.5 Conclusions

The conceptual framework proposed for this thesis is evaluated by a series of experiments, in this Chapter, we reported on the first experiment which utilised the use of an off-the-shelf tool, Synesketch, to extract emotions from the text. The tool provides a simple mechanism to identify six types of emotions in the text, based on the well established keyword technique. Synesketch showed high accuracy at the sentence-level, particularly when the sentence includes emotional words, and provides an emotion vector with accuracy around 83%. The experiment demonstrated the feasibility of the task, while of course highlighting the limitation of using a simple tool: for more complex sentences, with a complicated structure and figurative language, the use of the more advanced emotion/sentiment analysis algorithm will improve on the results.

The following Chapter will discuss the second experiment, when we report on how we can use the effect of extracted emotions in predicting the student's results.

Chapter 5

Experiment 2: Predicting Student's Final Grades

5.1 Introduction: Experiment 2 Overview

The work presented in this thesis is aimed at demonstrating the feasibility, as well as any technical challenges, of an "emotional observer" system, which would sit within a VLE and would be able to support educators in identifying potential situations at risk. Our conceptual framework, introduced in Figure 1, comprises a prediction system, which can decide to intervene, or simply flag, any observed cases for potential follow up by the lecturer. In the previous Chapter we tested the feasibility and limitations of using emotion analysis to label students' messages in private communication. In this Chapter, we concentrate on actual classroom behaviour, and we conduct a set of experiments, using well established data analysis techniques, in order to ascertain whether we can in fact find emotionally loaded text in classroom interactions, and up to which extent discovering emotionally loaded text is useful to predict students' performances.

Prediction and description are two primary objectives of data mining [90]. Prediction refers to the use of certain attributes or fields in the data set to forecast uncertain or potential values of other attributes of interest, with a process as shown in Figure [5.1]. Description refers to the process of discovering new patterns in the data, that humans can interpret.

In this Chapter we use data mining methods to pre-process and analyse students' data, and mine a students' performance prediction model, with the objective to understand any general relationships between different student' characteristics, including the emotions features and the likelihood to achieve good results.

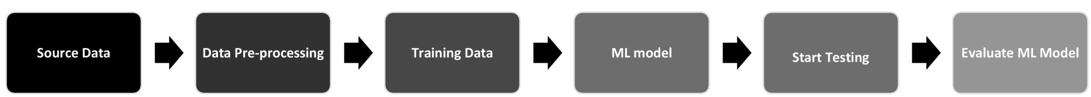


Figure 5.1: Structure of the prediction system

5.2 Prediction Tool: WEKA

As with Experiment 1, described in the previous Chapter, in the spirit not so much of finding the best possible prediction, but to simply demonstrate the feasibility of the task using off-the-shelf tools, we selected a tool called WEKA (Waikato Environment for Knowledge Analysis), version 3.8.2 [75, 81], a comprehensive tool covering all data mining steps, as one of the most commonly used open-source toolkits, comprising a collection of machine learning algorithms for solving data mining problems, and with the bonus of a simple and friendly Graphical User Interface (GUI). In this section we mention very briefly the main notions needed to understand our experiments, as they are not the focus of the study, and we refer the reader to any general textbook (for instance [83]) for a more comprehensive discussion.

5.2.1 Algorithm: Naive Bayes

The WEKA classifier algorithm used in this study is the Naive Bayes (NB) model [110]. A NB classifier is a simple model of probabilistic machine learning, used for classification tasks, built on Bayes theorem [88]:

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\ P(B) &= \sum_{Y} P(B|A)P(A) \end{aligned} \tag{5.1}$$

P(A|B) is the posterior probability calculated from P(A), P(B),and P(B|A), where

- P(A|B) is the posterior probability of class (A, target) given the predictor B: $B_1, B_2, \ldots, B_n$, where B_i is an attribute.
- P(A) is the prior probability of class, i.e. the class label.
- P(B|A) is the likelihood, i.e. the probability of the predictor P(B) given the class A.
- P(B) is the prior probability of the predictor.

The algorithm has many advantages:

- Lazy training: the algorithm does not need a large dataset for training because it depends on the property theorem, and can trained by little instance. It suggests that the algorithm can be effective with less training data as compared to other algorithms and that the capability of the algorithm is often a function of the amount of training data, whereas other algorithms may be unstable with insufficient data [34].

- Light weight in performance, so it is particularly useful for large datasets.

- Has a Java API available.

- Does not need as many parameters as Artificial neural networks (ANN) or Support Vector Machine (SVM) algorithms and similar.

The NB classifier has been used extensively to predict student grades. Mueen *et al* [140] conducted a study comparing three classifiers Naive Bayes, Multilayer Perception which is a most widely used form of Artificial Neural Network (ANN) that consists of multiple layers, and C4.5 (J48) which is used to envisage a Decision Tree (DT), and they found out that NB was the most accurate model among them, with overall prediction accuracy of 86%, for predicting student achievements (see Figure [5.2]). Similarly, a work by Devasia *et al* [48] found that NB could predict the performance of students at the beginning of the semester, outperforming other methods like Regression (which is refers to mathematical techniques that allow data scientists to predict a continuous value), Decision Tree (DT) (which is a flowchart represents a possible decision taken from all features), or Neural Networks (NN) which is refer to a group of algorithms are designed to imitate the human brain.

Classifier	Accuracy	Precision	Recall	Specificity
Naïve Bayes	86.0%	88.4	85.8	86.3
Multilayer Perception	82.7%	82.5	86.3	79.1
C4.5	79.2%	81.4	78.0	80.2

Figure 5.2: Comparison of classifiers on student grade prediction, reproduced from [140]

5.2.2 Measuring the Performances of Machine Learning

The performance of a learning algorithm when applied to data is measured by a metric called "Accuracy". This is calculated by using the formula:

- `Accuracy` $= \frac{\texttt{Number of samples Predicted Correctly}}{\texttt{Total Number of Samples}}$

$$\text{Accuracy} \ = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.2}$$

Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

Combined with accuracy, the other metrics used are Precision and Recall. Basically, precision gives a measure of how close are the predictions to one another (very distant predictions which average on a good accuracy are not a sign of a good outcome), while Recall provides an indication of missed positive predictions. Often a metric called F-Measure is used to provide a single statistic unifying precision and recall. While Specificity (True Negative Rate) refers to the ability of the classifier to distinguish negative results. The formulas for calculating these are as follows:

- `Precision = (TruePositives) / (TruePositives + FalsePositives)`

$$\text{Precision} \ = \frac{TP}{TP + FP} \tag{5.3}$$

- `Recall = (TruePositives) / (TruePositives + FalseNegatives)`

$$\text{Recall} \ = \frac{TP}{TP + FN} \tag{5.4}$$

- `F-Measure = (2 * Precision * Recall) / (Precision + Recall)`

$$\text{F-Measure} \ = \frac{(2 \cdot \text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}} \tag{5.5}$$

- `Specificity= (TrueNegatives) / (TrueNegative + FalsePositives)`

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{5.6}$$

5.2.3 Cross-Validation Method

When learning from a small data sample, one method for validating the result is to use *cross validation*: this procedure is a practical way to check how the algorithm classifies unseen data, and consists in splitting the data sample into a number K of groups, or "folds", and then taking in turn one fold out of the sample, training the algorithm on the remaining $K-1$ groups, and see how it performs when applied to the group that has been taken out. This reduces the variance and stabilises the accuracy by averaging over k different data subsets [167]. Most supervised learning algorithms, and WEKA as well, follow this method [81] It has been shown that generally a value $K = 10$ provides a model with low bias and variance [83]. We will describe the parameters for evaluating the results of this process when we show the experiments results in Section 5.4.

5.3 Dataset for Experiment 2: Stanford MOOCs

For this experiment we needed a dataset coming from actual class interactions, and we have opted for the Stanford MOOCs dataset, as made available by the Center for Advanced Research through Online Learning (CAROL) [35], and which has been used in a number of studies on learning analytics, e.g. [139, 150]. MOOCs, or Massive Open Online Courses, have become popular since the early 2010 as learning offered by a number of universities and training centres around the globe, offering often free courses in all disciplines and areas. Because of their nature, MOOCs see a huge dropout rate, and have therefore become an invaluable source to understand why students withdraw from online courses.

The Stanford MOOC dataset consists of data coming from courses taught by Stanford University and offered on various platforms (NovoEd[1], Coursera[2]and a Stanford installation of OpenEdX[3]. Data is anonymised and divided into a set of files extracted by OpenEdX and provided as Comma-Separated Values (.csv) files. These are listed, with the description provided by [35], in Table [5.1]. Attribute **anon_screen_name** is the anonymised participant identifier and serves as key to identify data related to the same student.

Table	**Content**
EventXtract	A much slimmed view of the OpenEdX tracking log events. The view only includes fields that are currently in use by the platform
ActivityGrade	Assignment grades; includes right/wrong for each problem part, the learners' solution choice for each answer part, and the first and last solution submission times.
VideoInteraction	An excerpt from tracking logs, focusing on participants' interactions with video.
Performance	Daily cumulative assignment performance per learner.
Demographics	Combines learner self-reported demographic information across multiple tables.
FinalGrade	FinalGrade contains the learners' grades as computed by the platform at the end of the course.
ABExperiment	Data resulting from the use of the OpenEdX AB testing facilities
OpenAssessment	Data resulting from the use of the OpenEdX peer grading facilities.
UserCountry	country of origin by IP address
EdxForum.contents	Forum entries since about June 2013. The entries are in relational form. Two versions can be generated: anonymised with hashed user ids different from the hashes used in other tables, and anonymized using the same hashes.
EdxQualtrics.question	Survey questions administered via the Qualtrics service.
EdxQualtrics.choice	Choices for multiple choice questions above.
EdxQualtrics.Response	Learner responses to survey questions.
EdxQualtrics.ResponseMetadata	Information about the learner behind each survey response. This table can be used to link survey responses to learners in the course.
CourseInfo	Facts about courses, such as start/end dates, academic year and quarter in which a course was offered.
EdxProblem	Metadata on problems offered to learners in courses.
EdxVideo	Metadata on videos in courses.

Table 5.1: Stanford MOOCs dataset files and description, from [35]

[1]`https://www.novoed.com/`
[2]`https://www.coursera.org/`
[3]`https://open.edx.org/`

5.3.1 Dataset Preparation

For our experiment, we needed to identify students for whom we could collect a complete set of data, and we concentrate on the following tables:

- `ActivityGrade` containing grades, including right/wrong answer and student's choice, and submission times;
- `FinalGrade` containing the grades computed at the end of the course;
- `EdxForum.contents` containing text entries to the discussion forums;
- `Alldata` an export table collecting engagement data, containing the name of the Name of the course, and other statistics per anon_screen_name, such as the length of each session.

From all the items in the dataset, we extracted 229 records, related to students who could be followed through all data files, through key attribute **anon_screen_name** cross-referenced with **student_id**, used in a group-by SQL query to collect a student profile, forming one vector for each student. For practicality, we classified the final grade into four classes: **Fail** under 50%, **Pass** 50-59%, **Merit** 60-69% and **Distinction** 70% over.

5.4 Experiment 2: Using WEKA on the Stanford MOOCs Dataset

This experiment focusses on the feasibility of producing a model which can predict students' performances, using data analysis, and including also emotionally loaded data. The latter characteristic of the experiment is what makes it different than the usual learning analytics experiments on student engagement, and it is in line with our conceptual framework, which proposes to design an "emotional observer" for the student activity. We divided our experiment into four phases. The first phase is a simple application of the NB classifier to the identified tables in the Stanford MOOCs dataset, without considering any emotion or sentiment, to understand what can be predicted from the data on the basis of "classic" learning analytics. Then, we consider only the portion of data for which we can perform an emotion analysis from text, that is data from traceable students who provided textual comments during the interaction with the course. On such subset, we first replicate the first phase, to create a baseline for comparison, and then we add the emotion element to the analysis. Finally, we analyse the impact of each emotion towards the prediction goal.

5.4.1 Phase 1: Testing the Dataset without Emotion Analysis

For the first phase in our experiment, from all the items in the dataset, we extracted 229 records, related to students who could be followed through all data files. Features are selected from three of the four tables we identified, that is **(ActivityGrade, FinalGrade, Alldata)**, so we exclude table **EdxForum.contents** containing text entries to the discussion forum. The features ***Percent_ grade, SessionLength(sec), NumEventsInSession, Module_ type***, and ***FinalGrade*** are selected to train the NB classifier. A sample of the content of these record is shown in Table [5.2] (record contents are included in Table [B.1] in Appendix B). Figure [5.3] shows the WEKA interface for the attributes to be trained by the tool.

Percent_grade	**SessionLength(sec)**	**NumEventsInSession**	**Module_type**	**FinalGrade**
100	13283456	235264	problem	Distinction
100	11873125	709544	problem	Distinction
33.33	30002	322	chapter	Fail
33.33	4944	12	problem	Fail
33.33	391923	14910	course	Fail
50	112800	975	chapter	Fail

Percent_grade	**SessionLength(sec)**	**NumEventsInSession**	**Module_type**	**FinalGrade**
33.33	71288	672	problem	Fail
66.67	550671	5973	chapter	Fail
50	480792	3000	problem	Fail
83.33	1964286	15225	chapter	Fail
100	1552680	12000	course	Merit
100	283275	900	sequential	Merit
100	3355380	66780	problem	Pass
100	398060	1404	problem	Pass

Table 5.2: Features used in Phase 1 for the classifier

Table [5.3] shows the accuracy result for using the NB classifier for predicting student's *FinalGrade* from the features selected, when applied to the identified 229 records, and using a cross-validation with k-fold 10. The correctly classified instances were 185, the incorrectly classified instances were 44. The output from the WEKA interface is shown in Figure [5.4] and shows the results for each of the grade classes (Distinction, Merit, Pass, Fail).

Correctly Classified Instances %Accuracy	**Incorrectly Classified Instances %Accuracy**	**Precision**	**Recall**	**F-Measure**
80.78	19.21	0.755	0.808	0.777

Table 5.3: Naive Bayes algorithm accuracy for Phase 1

In the next experiment phase 5.4.3, we will use other features for more reliable results in an educational platform.

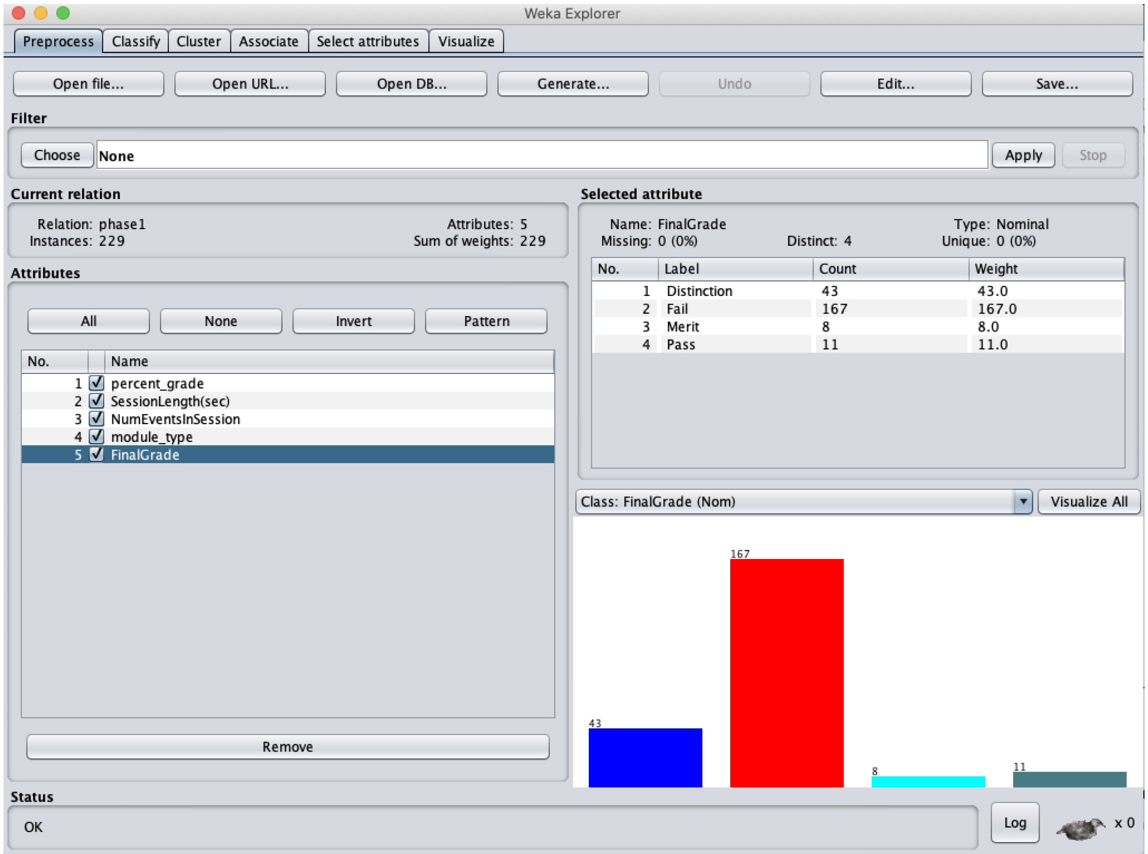


Figure 5.3: WEKA interface - Features used for training in Phase 1

5.4.2 Phase 2: Focus on Records with Textual Comments

In this phase, we extracted, from the 229 records identified in the first phase, those for which there were correlated entries in the `EdxForum.contents` table of textual comments. This gave us a subset of 71 records. In this phase, we repeated the same experiment in Phase 1, but restricting the dataset to these 71 records. We note that among these records, no student received a final **Merit or Pass** grade, so we test only for **Fail** and **Distinction** (see Fig. 5.5).

We train the NB classifier as in the first phase, by using features ***SessionLength(sec),***

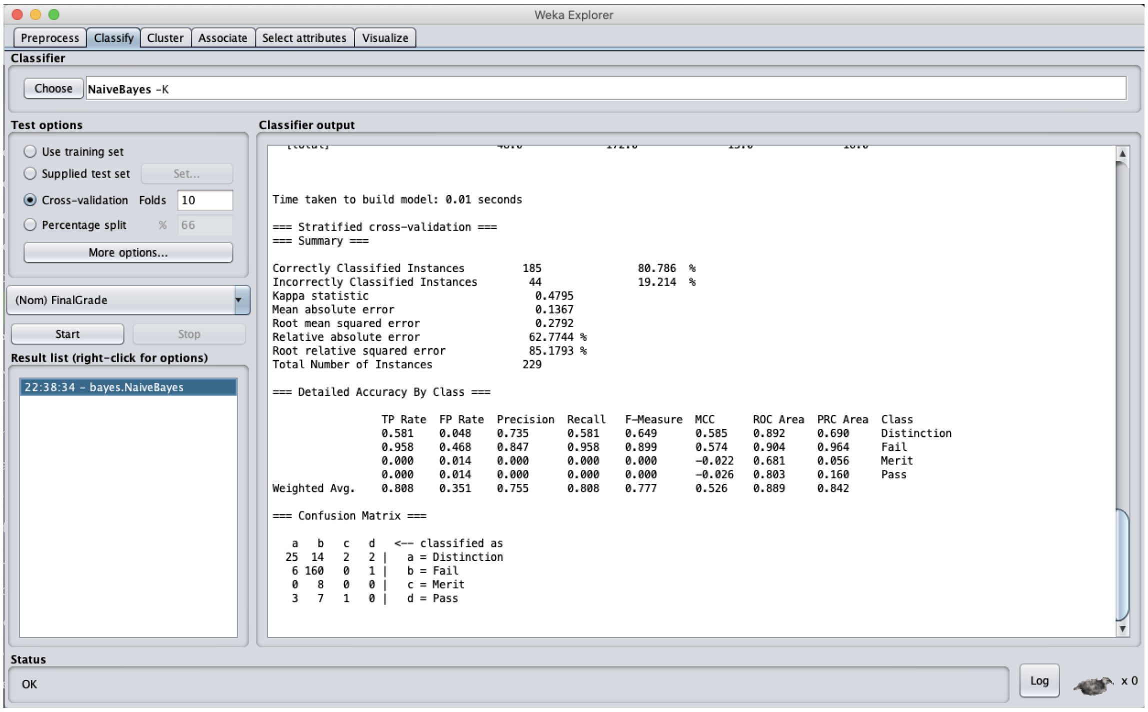


Figure 5.4: WEKA result of the NB classifier in Phase 1

NumEventsInSession, percent_ grade, down_ count, up_ count, and ***FinalGrade***, Where we add down_count (Number of down votes) and up_count (Number of up votes) attributes from the `EdxForum.contents` table. Then, we perform a 10-fold cross validation. This time the correctly classified instances were 91.5%. The summary of the statistics is given in Table [5.4], and the output from the tool is in Figure [5.6].

Correctly Classified Instances %Accuracy	**Incorrectly Classified Instances %Accuracy**	**Precision**	**Recall**	**F-Measure**
91.54	8.45	0.918	0.915	0.916

Table 5.4: Naive Bayes algorithm accuracy for Phase 2

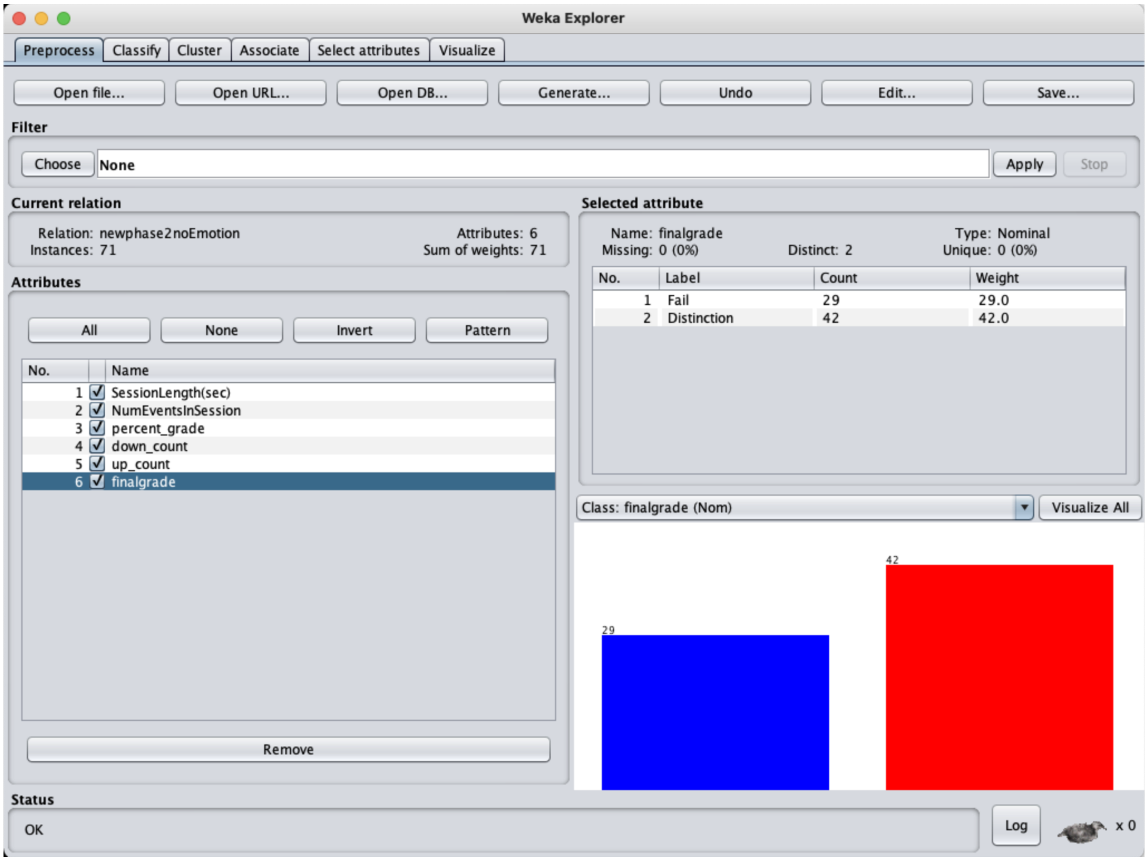


Figure 5.5: Features used in Phase 2

5.4.3 Phase 3: Prediction adding Emotion Analysis

In this phase we use again the 71 instances, extracted from the initial 229 records, and we add to the training the textual comments associated to these records from table `EdxForum.contents`. In order to prepare the textual comments for use by the classifier, we employed the emotion analysis tool used in Experiment 1 (Section 4.2), Synesketch, to perform a sentence analysis of the comments. We noted in Experiment 1 that the tool performs reasonably well with short sentences, as these are, and we extracted from the comments features corresponding to the six emotions recognised ((*Anger, Disgust, Fear,*

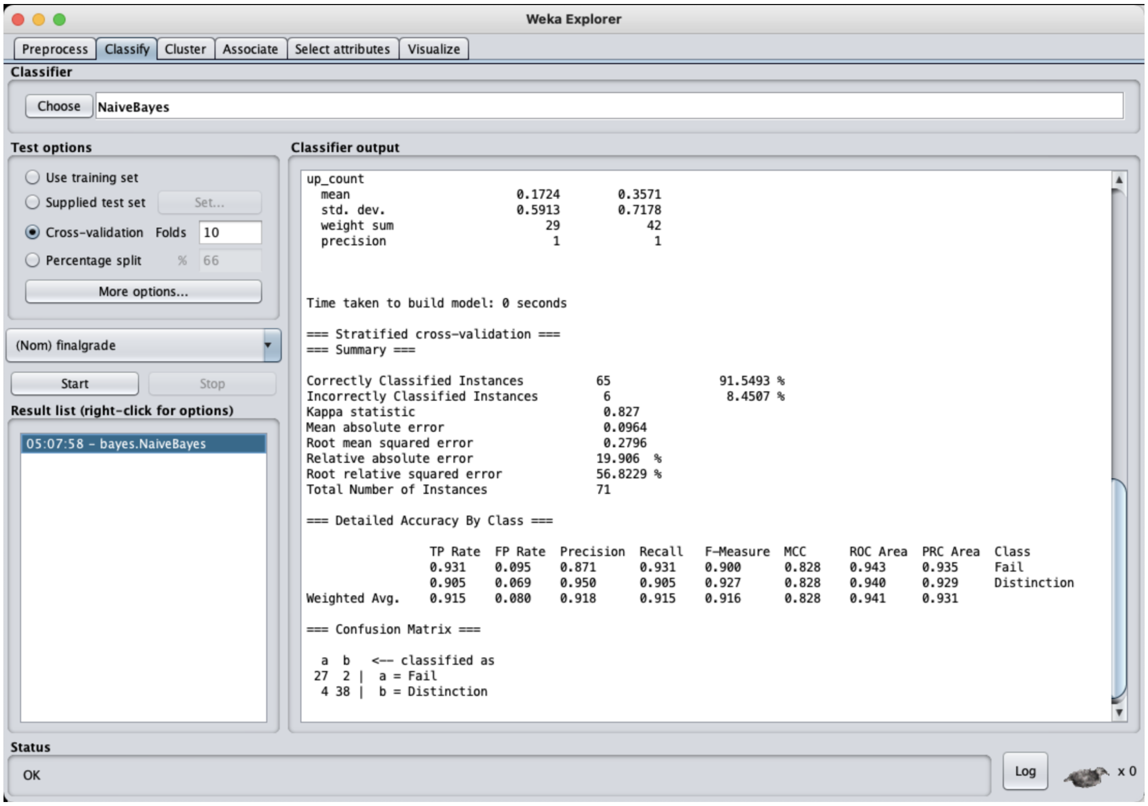


Figure 5.6: Accuracy result of Naive Bayes to predict final grade without emotion features

Happiness, Sadness,Surprise). Table [5.5] shows a sample of this new dataset, while a complete set is in Table [B.2], in Appendix B. Figure [5.7] shows the WEKA interface for the attributes to be trained by the tool.

With this augmented dataset, including all features used in the previous two phases, plus the emotions identified, we again applied the NB classifier, applying a 10-fold cross validation. The output from the tool when adding the emotion features is shown in Figure [5.8], and the summary of the statistics is Table [5.6].

SessionLength(sec)	NumEventsInSession	percent_grade	down_count	up_count	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Finalgrade
43434	342	100	0	0	0	0	0	0	0	0	Fail
112800	975	50	0	0	0.1	0.1	0.1	0	0.1	0.1	Fail
119102	1240	50	0	0	0	0	0	0	0	0	Fail
2629604	36076	100	0	0	0	0	0	1	0	0	Distinction
2629604	36076	100	0	0	0.2	0	0	0.1	0.1	0.2	Distinction
2629604	36076	100	0	0	0	0	0.1	0.3	0.3	0	Distinction
3216918	28290	25	0	0	0	0	0	0.1	0.1	0	Fail
3816736	81543	100	0	0	0	0	0	0	0	0	Fail
3816736	81543	100	0	0	0.7	0.7	0.8	0.1	0.1	0	Fail
4873033	58916	100	0	0	0.7	0.7	0.8	0.2	0.3	0.1	Distinction
7570450	224480	83.33	0	3	0.1	0.1	0	0.1	0.1	0.1	Fail
7570450	224480	83.33	0	0	0.1	0.1	0.1	0.3	0.1	0.1	Fail
10445416	150388	100	0	3	0.1	0.1	0.2	0.1	0.1	0	Distinction
10445416	150388	100	0	0	0	0	0	0.1	0	0	Distinction
10445416	150388	100	0	2	0.1	0	0	0.3	0.3	0.1	Distinction

Table 5.5: Adding emotion features to the dataset in Phase_3

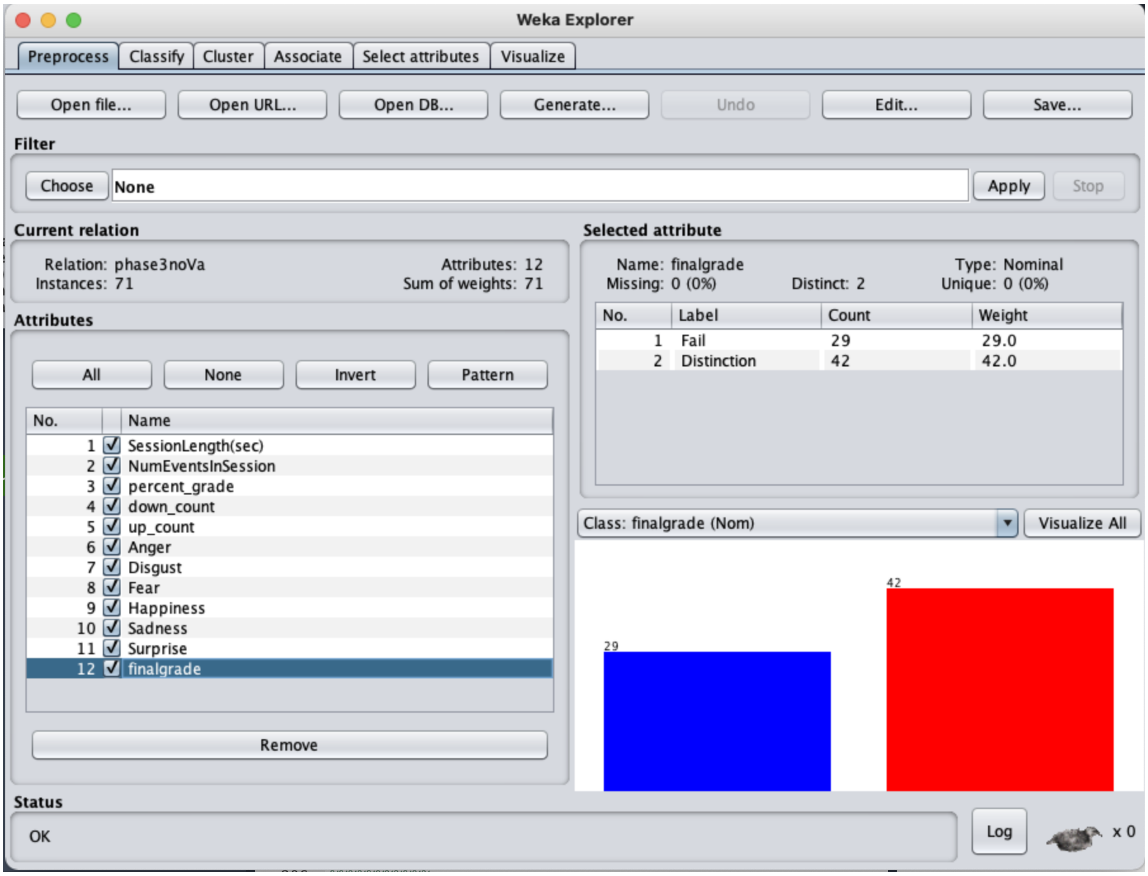


Figure 5.7: Features used in Phase 3

Correctly Classified Instances %Accuracy	**Incorrectly Classified Instances %Accuracy**	**Precision**	**Recall**	**F-Measure**
88.73	11.26	0.887	0.887	0.887

Table 5.6: Naive Bayes algorithm accuracy for Phase 3

5.4.4 Phase 4: Most Influential Emotions

If we compare the results of Phases 1 till 3 (see Table [5.7]) we do not seem to arrive to a conclusive result: this is not unexpected, in the sense that much more work need to be done should these phases be implemented for real, both from the point of view of testing

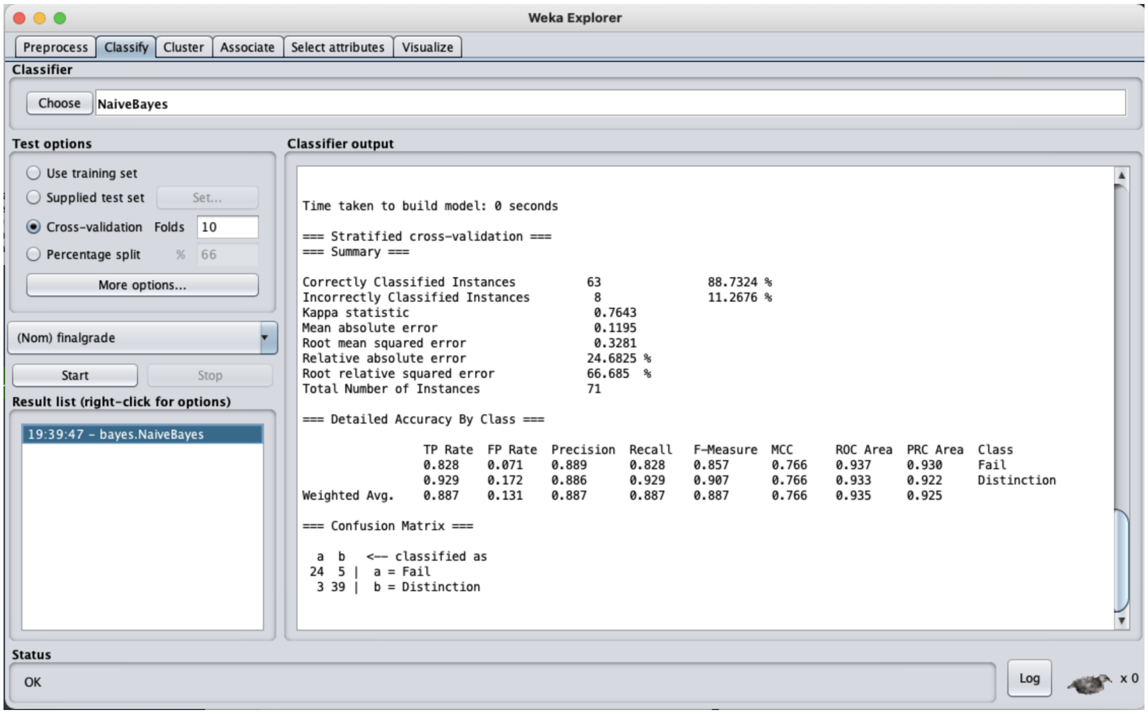


Figure 5.8: Accuracy result of Naive Bayes to predict final grade with emotion features

different classifiers, and from the point of view of selecting a more sophisticated emotion analysis tool, as we discussed in the previous Chapter. Nevertheless, we want to push our analysis further, in the spirit of our feasibility test, and see what can be done if we decide for instance to have a closer look at the different student emotions, and how each of these individually can influence the prediction.

This task is also common in data analysis: when performing a classification using machine learning, if the number of features increases, not only the training time but also the risk of *over-fitting* increases exponentially [25, 227], where over-fitting is a situation where the model is so complex that it "learns the noise" rather than actual patterns in data. A way to prevent this is to use "Feature selection", where less relevant features are excluded from the model to make it more accurate and reliable [25, 48, 130, 227]. Feature

	Accuracy (TP)%	Precision	Recall	F-Measure
Phase 1				
Distinction	58.1	0.735	0.581	0.649
Fail	95.8	0.847	0.958	0.899
Weighted Average	80.8	0.755	0.808	0.777
Phase 2				
Distinction	90.5	0.950	0.905	0.927
Fail	93.1	0.871	0.931	0.900
Weighted Average	91.5	0.918	0.915	0.916
Phase 3				
Distinction	92.9	0.886	0.929	0.857
Fail	82.8	0.889	0.828	0.857
Weighted Average	88.7	0.887	0.887	0.887

Table 5.7: Comparison of the NB performances for the three phases

selection consists of an evaluation of attributes, to determine which belong to the context of the output and which do not, and a search method, where various attribute combinations are explored to come up with a shortlist of the selected features [130, 172].

The most widely used technique for or selecting the most relevant feature in a dataset is *correlation*, based on Pearson's correlation coefficient [3, 52], and involved calculating this coefficient between each attribute and the output variable, then selecting only the attributes with a moderate to high positive or negative (coefficient close to $+1$ or to -1) and removing all low-correlation attributes (coefficient close to 0) [180].

WEKA supports the selection of correlation-based features using the ***CorrelationAttributeEval*** feature, which involves the use of a Ranker search tool, evaluating each attribute and listing the results in a ranked order. Figures [5.9] and [5.10] show the configuration setting in WEKA for attributes CorrelationAttributeEval and Ranker.

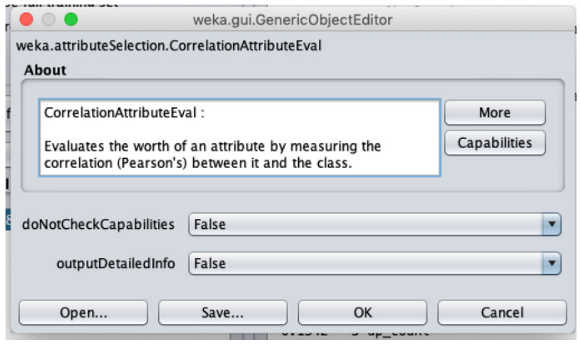


Figure 5.9: Configure CorrelationAttributeEval

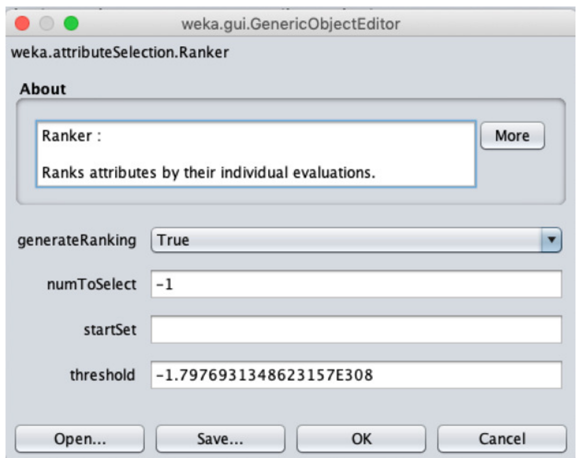


Figure 5.10: Configure Ranker

Figure [5.11] shows the result of running the Features Selection process on our dataset, with the suggestion that one attribute, Session Length, has the highest correlation with the output class, and provides a rank of the remaining attributes. If we set a coefficient cut-off for relevant attributes equal to 0.1, then the remaining attributes *(Disgust, Fear, Surprise, down_ count)* could possibly be removed to improve the performances.

Figure [5.12] shows the result of applying the classifier to the dataset of Phase 3 after removing such non-relevant attributes: the new accuracy, result of feature selection is (91.54%) which is 3% higher than with all emotion features which was 88.73%. The summarised output is given in Table [5.8].

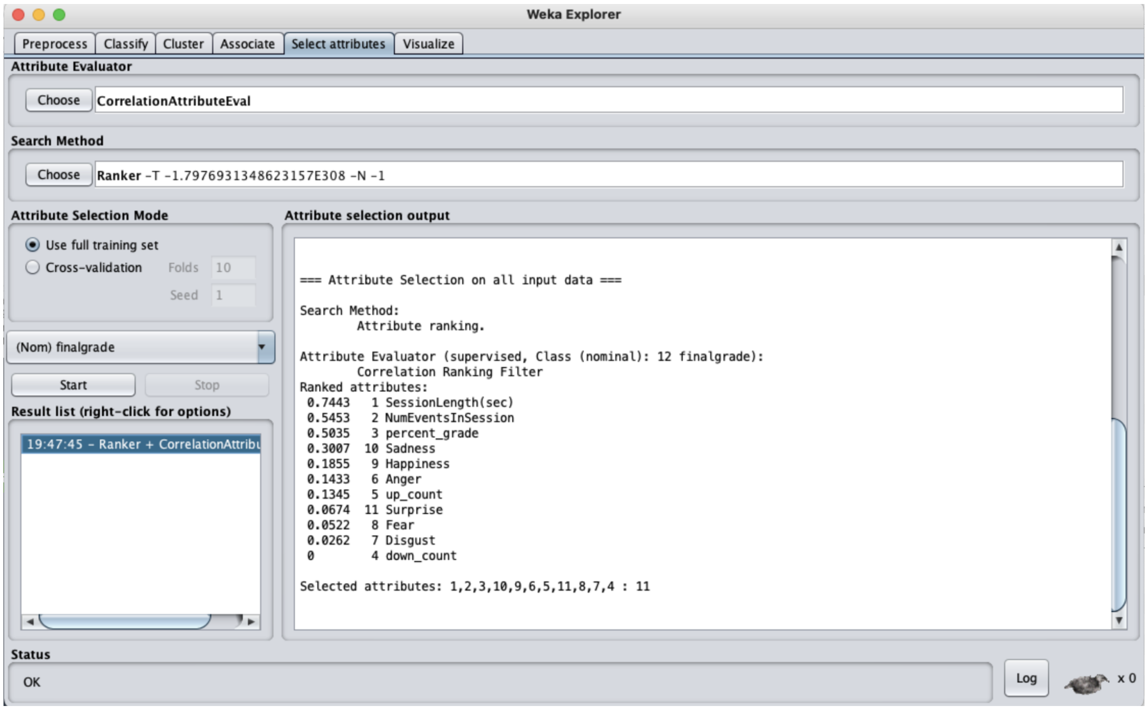


Figure 5.11: Feature Selection result for all attributes

Correctly Classified Instances %Accuracy	**Incorrectly Classified Instances %Accuracy**	**Precision**	**Recall**	**F-Measure**
91.54	8.45	0.918	0.915	0.916

Table 5.8: NB classifier accuracy after removing non-relevant attributes

5.4.4.1 Analysing the Impact of each single emotion

As a final part of the experiments, we performed a study by selecting each emotion in turn and considering only that emotion in addition to the other features for our NB classifier,

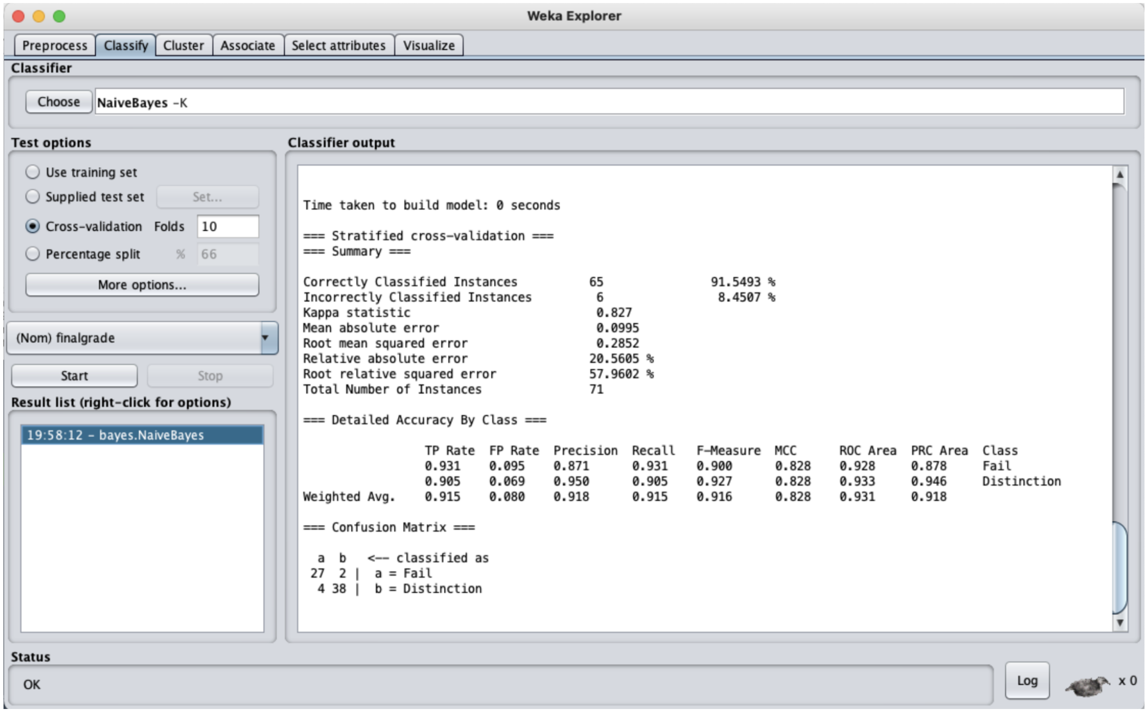


Figure 5.12: NB classifier accuracy after removing non-relevant attributes

using therefore as dataset the one used in Phase 3 but with one single emotion in turn (see extract in Table [5.9] for when the selected emotion is Anger). The average results when adding each emotion in turn is shown in Table [5.10]. Sadness was the most influential emotion to predict final grade, and this is consistent with the previous result of feature selection shown in Figure [5.11], where sadness was the top ranked attribute in terms of relevance to the final grade. In fact, the model with the inclusion of the Sadness feature only outperform every other model used in Phase 1, 2, or 3. The statistical significance of the difference between these values have been assessed through a descriptive statistics, analysis of variance (One-way ANOVA), in order to identify the most influential emotion, and the summary of the ANOVA test between the variables showed in the table [5.11], (complete analysis between the different pairs are included in Appendix B section [B.3].)

SessionLength(sec)	NumEventsInSession	percent_grade	down_count	up_count	Anger	Finalgrade
43434	342	100	0	0	0	Fail
112800	975	50	0	0	0.1	Fail
119102	1240	50	0	0	0	Fail
119102	1240	50	0	1	0.1	Fail
958783	9457	100	0	0	0.1	Fail
1058080	17320	100	0	0	0.1	Fail
1623812	54199	100	0	0	0	Fail
1623812	54199	100	0	0	0	Fail
1623812	54199	100	0	0	0.1	Fail

Table 5.9: Part of the dataset containing only one emotion feature

Emotion Feature	Correctly Classified Instances %Accuracy	Incorrectly Classified Instances %Accuracy	Precision	Recall	F-Measure
Anger	91.54	8.45	0.918	0.915	0.916
Disgust	91.54	8.45	0.915	0.915	0.915
Fear	91.54	8.45	0.915	0.915	0.915
Happiness	90.14	9.85	0.902	0.901	0.902
Sadness	92.95	7.04	0.930	0.930	0.930
Surprise	90.14	9.85	0.902	0.901	0.902

Table 5.10: Model accuracy with each emotion

Groups	Count	Sum	Average	Variance		
Happiness	71	15.348394	0.21617456	0.08342298		
Disgust	71	4.796696	0.0675591	0.02503918		
Fear	71	5.15031	0.07253959	0.0228789		
Anger	71	5.64398	0.07949268	0.02008261		
Sadness	71	10.384967	0.14626714	0.0206287		
Surprise	71	3.153322	0.04379614	0.01607516		
Source of Variation	**SS**	**df**	**MS**	**F**	**P-value**	**F crit**
Between Groups	1.48839739	5	0.29767948	9.5049707	1.32901E-08	2.23542594
Within Groups	13.1850023	421	0.0313183			
Total	14.6733997	426				

Table 5.11: Summary of the One-way ANOVA test for feature set comparisons

5.5 Conclusions

In this Chapter we performed a feasibility study for the portion of our conceptual framework concerning the analysis of student performance in the classroom and how this may or may not correlate with the emotions the students are expressing through their interactions. The experiments made use of an off-the-shelf data analysis tool, WEKA, and were applied to an open source dataset based on Stanford MOOCs.

The limitation of the specific experiment are of course evident, starting from the technical point of view: just as we discussed in the previous Chapter on the suitability of Synesketch to the task, here more fitting algorithms need to be explored for machine learning, with a dedicated study to decide the most appropriate classifier to the task at hand. Nevertheless, the experiments demonstrated the technical feasibility of the task, and were crucial to collect useful considerations towards a possible implementation.

An impactful limitation to mention here comes from dataset itself, which, by coming from a MOOC type of educational offer, is not expected to show much interaction among participants in the module: more traditional environments, in which the same cohort of students sits several modules in their progression towards obtaining a formal degree will create more cohesion in the classroom, and therefore a higher number of interactions: the textual comments in the Stanford dataset were simply not enough to make a difference.

This brings us to the next step of the study, which relates precisely to this issue, i.e. that our framework is intended to be applied to traditional online situations, where the notion of cohort exists. The next Chapter will explore this notion, and will complete our discussion around the feasibility of the framework.

Chapter 6

Experiment 3: Emotional Profile and the Notion of Cohort

6.1 Introduction: Experiment 3 Overview

In the previous Chapters we have explored some of the tasks that a system implementing our conceptual framework would need to tackle, and we have demonstrated their technical feasibility by using simple, off-the-shelf tools, tested on available datasets. However, the tasks considered were not too far from what a complex learning analytic tool might do, and are far from the more sophisticated concept of "emotional observer" we have included in our framework. In this Chapter we want to explore this further, and consider the complexity of putting in place such a notion.

The main limitation of the experiments seen so far, and of the use of learning analytics in general, even with the added emotion element, is that each episode/student is considered in isolation: a tool would detect a situation, and create a trigger, but the situation is seen as an independent event. However, real life scenarios are much more complex, and from an educator perspective it is not very informative to be able to detect that a student's communication is of a certain type, or has a certain emotional content: it could be that this student has always communicated in that particular way, and the aggregate of the messages captured in one course is just perfectly in line with the usual behaviour. In Experiment 2 we have seen that the correlation between emotional content and grades can be quite loose: students could be classed as overall "sad" and still get good grades in the end.

Also, the communication from a student cannot be considered in isolation without knowing what is going on in the classroom: has the course get to a point where a complex topic is being discussed? Is there a deadline approaching? Is there, for courses taken in parallel, an event in *another* classroom that is impacting the behaviour of this class? Is the specific student cohort of this particular classroom especially edgy, for one reason or another? These are all situations that an educator knows all too well, and that we would like to be able to capture in an ideal "emotional observer" system.

ln this Chapter we want to explore the feasibility of two tasks:

1. Creating a baseline for a student emotional profile, by looking at the textual output of a student. This baseline will incorporate the notion of "style" of the student's text, in terms not only of emotional content, but also on writing features, such as how rich the text produced by the student is, and what characteristics the Part of Speech (POS) of this student's usual textual output show. This will help establish whether a student's writing style is out of line with the norm for that particular student, and therefore whether this is a situation which may need to be explored further by the educator.

2. Creating a notion of classroom, or cohort, to be able to establish whether the textual output of a particular student is out of line when compared with the rest of the classroom, and therefore, again, whether this situation needs to be explored further.

The main aim is, we repeat, not to create a system that could substitute the teacher, and intervene, but to create a mechanism to alert the teacher that something might need to be looked at. In summary, we envisage a flow of events like the one in Figure 6.1.

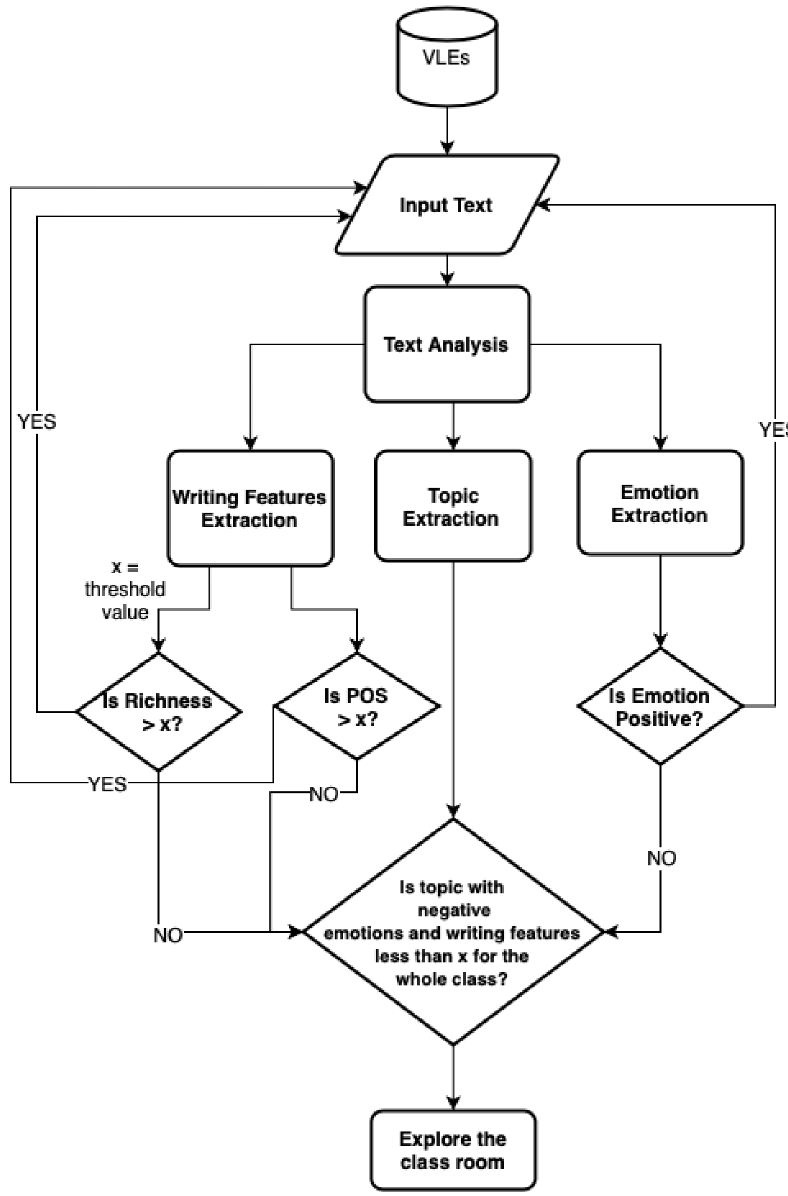


Figure 6.1: The emotional observer flowchart

Testing the feasibility of such system requires a more sophisticated text analysis algorithm, and more importantly, a more suitable dataset, which is very different from those made available by online education providers, and in this Chapter we detail how we approached the problem.

6.2 Text Analysis Tool: Algorithms for Stylometric Analysis

Stylometry traditionally is the field of research which studies the characteristics of "style" (usually of text, but also music or art) [119], usually with the aim of determining authorship, and has now found a new popularity thanks for the massive increase in the availability of text over the Internet. Stylometric analysis, as a statistical analysis of textual input, has been used at example for studying blogging and micro-blogging (Twitter) [21, 71] or also to improve cybersecurity algorithms [197].

For Experiment 3, we perform a word level classification, devising an algorithm based on previous work by Whissell [214], who used stylometric analysis to study the evolution of the lyrics of Beatles' songs over the years, using the *emotion clock* where categorical emotions are based on Russell's model dimensions [171], so for instance establishing how John Lennon's lyrics become less and less cheerful over time. This is the type of task we need to perform on our students' textual input, in order to determine whether something has changed in the style of writing throughout the studies from enrolment. Whissel demonstrated that whole passages of text can be associated with a measure quantifying the emotional meaning of the words forming them, and the addition of emotion improves on other classifications based on classical stylometric measures like length of text, use of pronouns etc.

The algorithm calculates the Emotion Polarity, based on the frequency of emotional word during various stages. We do this for 12 emotional states, using the emotion clock [214]: *(peaceful, friendly, admiring, cheerful, brave, alert, alarmed, furious, distant, sad, depressed, indifferent).*

The algorithm works as follows (and is depicted in Figure [6.2]):

Algorithm 2 Calculating the Emotion Value

1. Initialise an Array of words associated to the 12 emotions: this is done using dictionary.com and including all associated words in the definition of each emotion, all synonyms and associated words;
2. For each text "session", that is a portion of text output by a user (Client010_a in the picture):
   (a) parse the session using WordNet® for reference, and collect all words and all synonyms for such words
   (b) compare set of words for the session with the emotion Array of words for each emotion
   (c) produce a vector containing the number of words in session for each emotion; note that this is not a vector of normalised values, as we want to represent also roughly how verbose this participant was when expressing the emotion, as a useful stylistic metric that can signal change in a person's attitude or mental state.

In addition to this value, we use the following two metrics to measure a passage:

- **Richness ratio:** this is a measure of the size of the writer's vocabulary, to calculate this we use the technique in [196], and we calculate the Total different words in the passage, normalised by total number of words (tokens). It has been demonstrated that a change in this statistic is a sign that writer has changed writing style [2, 196].
- **POS, or Part of Speech ratio:** this is a measure of the extent to which the writer

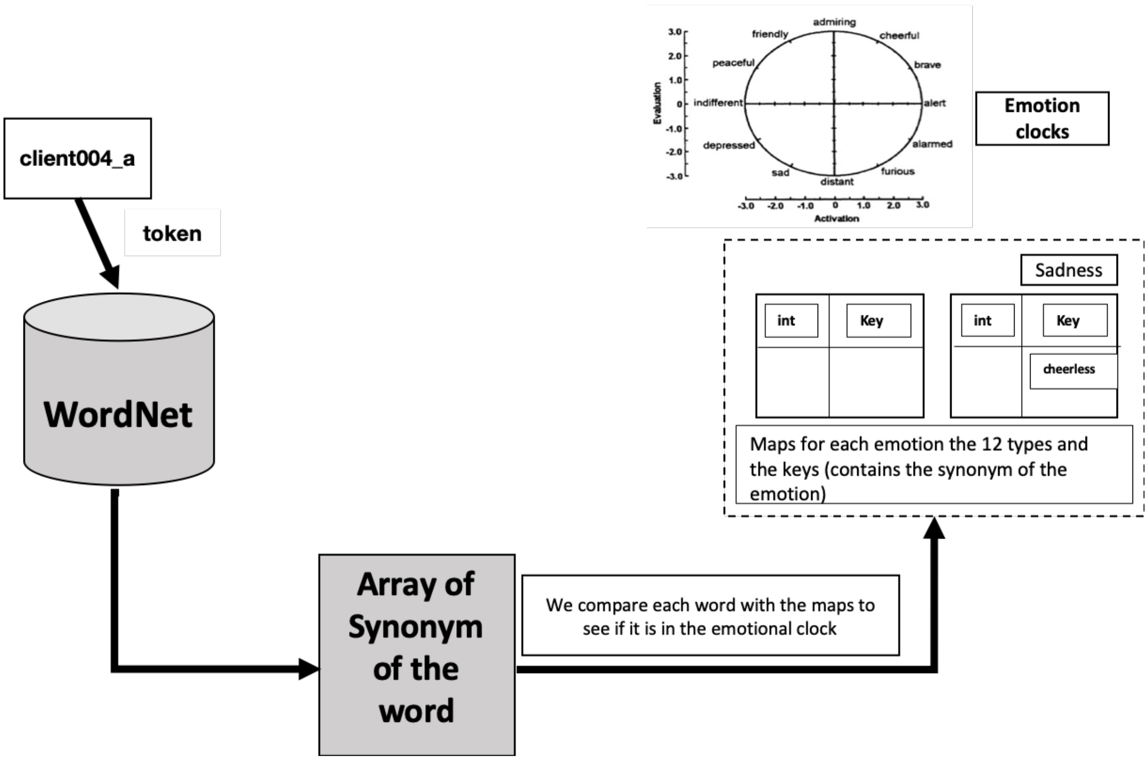


Figure 6.2: Emotion extractions and analysis

uses well formed sentences, as opposed to sentences that break or interrupt. This is calculated very coarsely by calculating the number of sentences which contain the three main parts of speech (nouns, verbs, adjectives) normalised by the total number of sentences in the passage (where a sentence boundary is determined by a full stop).

In the next section we describe how we put together a dataset in order to test the above measurements.

6.3 Dataset for Experiment 3: Motivational Interview Corpus

The educational resources we have explored in order to find a suitable dataset for this experiment did not provide much choice. In our requirements, we wanted to be able to capture a series of conversation over potentially a long period of time from the same student/user, and which could give us enough textual output to be able to perform a good stylometric analysis. We wanted to be able to follow the same student over a number of modules. We also wanted to analyse cohorts of students, to be able to capture the "emotional state of the classroom".

We found a suitable source when we moved outside of the educational boundaries, and in particular when we analysed datasets coming from psychology and behavioural sciences. In particular, the field of Motivational Interviewing [168] provided a good source of conversational data. Motivational Interviewing is an approach to counselling aimed at helping clients modify their behaviour by going through a dialogue which prompts questions inspiring positive change.

We use for this experiment a Motivational Interviewing corpus of transcripts [5]. The corpus consists of a searchable collection containing real transcripts of counselling and therapy sessions between a client and counsellor. The database contains more than 2,000 session transcripts, 44,000 pages of client narratives, and 25,000 pages of secondary reference material. Transcripts are anonymised, but grouped by sessions related to the same client, so it is possible to track the same client over a number of counselling sessions, and this was particularly important. For each anonymised client, the source provides information on age, gender, marital status, any symptoms or condition, as well as some general information on the therapist, such as gender and level of experience.

The index of all sessions on the database provides the client code and a brief description of the issue (see Table 6.1 for some examples).

For each client, the database provides a number of sessions, where it is also indicated a list of "subjects", the main issues under discussion during the session. There is a number of emotional state and personality factor which is annotated by the therapist. Also, it

Session Transcripts
Client 004: Heterosexual male between 25-30 years of age suffers from dysphoria, (1970), pp. 1-28
Client 011: Female in her early twenties suffers from a lack of self-confidence and problems in interpersonal relationships, particularly with men, (1971), pp. 1-39
Client 016: Female in her early twenties presents with a sense of lack of control in her life; she wants to be her own person, (1971), pp. 1-11
Client 019: Female in her early twenties presents with anxiety and lack of confidence, (1971), pp. 1-23
Client 031: Client presents in therapy with low self-esteem, (1971), pp. 1-10

Table 6.1: A sample of the session in transcripts corpus

is include information about the client demographic, gender and age (see an example in Fig. 6.3). For each session, a transcript of the conversation is provided (see example of an extract in Fig. 6.4).

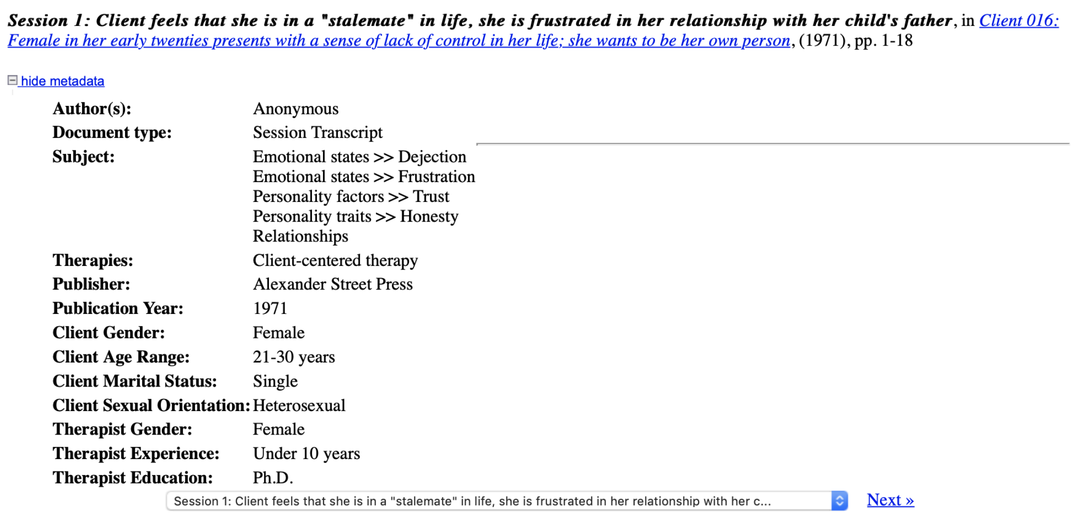


Figure 6.3: Sample of Client016 Session1 information

The choice of the corpus has some important benefits:

1. the conversations are very rich in terms of topics and emotional load, so the corpus is a useful testbed to understand how emotion and style of writing interact;

2. the conversations have been labelled by the curator of the corpus, so we have an objective way to identify parameters in the dialogues;

3. the conversations happen over more or less long periods of time, between the same

---

BEGIN TRANSCRIPT:

23016 - First Interview

COUNSELOR: Are you afraid that it will be too young by these standards?

PATIENT: No, I don't know I just feel like I've been talking [inaudible at 00:34].

COUNSELOR: I'm actually 28. Does that make you feel any better?

PATIENT: A little.

COUNSELOR: What is it that you have on your mind?

PATIENT: Well, did you see any of the tests that I took? Well, I guess basically it's I feel like I'm in a stalemate and I can't make up my mind for myself. I just can't make any decisions and stick to it, do you know what I mean? I can't seem to move forward.

COUNSELOR: It feels like you're trapped where you are now?

PATIENT: Not so much trapped as that I don't know what to do. I guess the thing is I don't trust myself. I don't feel secure in myself to depend on myself and do what I want. And sometimes I even feel confused about what I think.

https://asp6new.alexanderstreet.com/psyc/psyc.object.details.aspx?dorpID=1000061050 Page 1

COUNSELOR: Do you resent him?

PATIENT: [inaudible at 58:31]

COUNSELOR: Do you have parents in New York?

PATIENT: Well they're in Ithaca.

COUNSELOR: Do you see them?

PATIENT: Yeah, they get along real well with Adrian and I have a sister who is about three blocks from me.

COUNSELOR: She helps you out sometimes?

PATIENT: Yeah, they're watching him right now. My father hates Robert and won't talk about him. He's made friends with my sisters and my mum, but my father no way.

COUNSELOR: Do you feel like that's unfair?

PATIENT: No. Ever since I got pregnant he's just been... I guess basically he's been honest but he's also done it in poor taste. My father just was turned off by him and doesn't want to deal with him.

COUNSELOR: [inaudible at 59:41]

PATIENT: He doesn't know that [inaudible at 59:46]

COUNSELOR: And your mother does know him?

PATIENT: At first she did but I don't know how she feels right now.

COUNSELOR: You don't seem much pressured by them.

PATIENT: No. We've got four girls in our family. He's been getting it from all sides so he'll say things every once in a while but he wouldn't pressure me. [60:16]

END TRANSCRIPT

Figure 6.4: Sample of Client016 Session1 transcripts

client and the same counsellor, so they give us the opportunity to follow the same person over time, and to identify possible change in emotion or writing style.

However, the corpus has the following obvious drawbacks:

1. each conversation happens independently, all clients have individual sessions with the therapist, so there is no notion of "group", something we said we needed to consider, in order to analyse a "classroom" behaviour;

2. the domain of conversation is not education.

While there is nothing we can do with respect to the second drawback, which we have to accept as there was no equivalent dataset we could have used in the educational domain, we can work on the first drawback and create a simulated environment, which is the exercise we will present in this section.

6.3.1 Creating a Dataset from the Motivational Interview Corpus

For Experiment 3, we built a dataset, out of the transcripts of the Motivational Interview Corpus, by grouping and labelling the sessions in the way explained below.

We extracted data related to all clients in the corpus, excluding those for whom only one session was present: this gave 76 clients in total. We gathered together all sessions related to the same client, and this constitute the textual output of the client. For the purpose of the experiment, we considered only the client's turns in each dialogue, so we discarded the counsellor comments, and any other comments by the scribe (e.g. annotating pauses etc).

For each session *Session_name*, we created a *Vector={Session_name, Topic, Emotion, Emo_value, Richness, POS}* where each field has the following meaning:

1. Topic = the general topic title for the session, as retrieved from the transcripts);

2. Emotion = the most predominant emotion of the session (emotion type), calculated with Algorithm 2

3. Emo_value = the value, or size of the predominant emotion: calculated in number of words;

4. Richness = richness ratio, calculated as in the formula previously presented;

5. POS = part of speech ratio, again calculated as in the formula previously presented.

At the end we export all results into a comma-separated values (.csv) file. An example of this file is in Table 6.2, while Table [C.1] in Appendix C contains the complete set for all the 76 clients' sessions.

Note that there are other features that the transcripts report that we chose not to use, most notably the Emotional states, and the Personality factors and traits (see Fig. 6.3). While this information would be very useful to evaluate an emotion extraction algorithm, for the purpose of this particular experiment we wanted to rely exclusively on the stylometric features, even if in a simplified form: we would not think a realistic scenario where the students receive an expert psychological evaluation for each course they take on.

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Ability	Client004_a	sad	20	0.334792123	0.411255411
Relationships	Client004_b	sad	40	0.333333333	0.495934959
Relationships	Client004_c	sad	33	0.310679612	0.506410256
Relationships	Client004_d	sad	22	0.426035503	0.41509434
Relationships	Client004_e	sad	9	0.381322957	0.397590361
Behaviour	Client004_f	sad	15	0.388779528	0.352601156
Behaviour	Client004_g	peaceful	21	0.359242325	0.473170732
Behaviour	Client004_h	indifferent	11	0.355733662	0.497409326
Ability	Client004_i	sad	177	0.200187091	0.450715421
Ability	Client011_a	friendly	945	0.112697198	0.824940048

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Ability	Client011_b	friendly	146	0.229634672	0.83203125
Relationships	Client011_c	friendly	64	0.29015919	0.833333333
Relationships	Client011_d	friendly	126	0.260787992	0.884615385
Relationships	Client011_e	friendly	143	0.256395178	0.752941176
Ability	Client011_f	peaceful	67	0.238571815	0.90797546
Behaviour	Client011_g	peaceful	54	0.199496855	0.79342723
Culture	Client011_h	friendly	113	0.22172619	0.790598291
Relationships	Client011_i	friendly	142	0.265526553	0.820512821
Relationships	Client016_a	peaceful	41	0.315902579	0.669421488
Relationships	Client016_b	friendly	14	0.433849821	0.627118644
Relationships	Client016_c	friendly	46	0.343678686	0.645714286
Relationships	Client016_d	peaceful	49	0.382887189	0.725352113
Relationships	Client016_e	peaceful	41	0.415368082	0.687022901
Development	Client016_f	peaceful	33	0.364278507	0.78030303
Relationships	Client016_g	peaceful	36	0.278882576	0.74796748
Behaviour	Client016_h	peaceful	29	0.322977346	0.666666667
Behaviour	Client016_i	peaceful	32	0.304277206	0.833333333
Development	Client018_a	sad	73	0.248955224	0.699152542
Relationships	Client018_b	peaceful	51	0.302836596	0.716666667
Behaviour	Client018_c	sad	57	0.320996979	0.75
Ability	Client018_d	sad	39	0.292576419	0.668639053
Personality	Client018_e	peaceful	63	0.274716029	0.624277457
Relationships	Client018_f	sad	78	0.301617149	0.71
Personality	Client018_g	sad	42	0.361516035	0.823529412
Relationships	Client018_h	peaceful	59	0.241905471	0.628318584
Ability	Client018_i	peaceful	71	0.344947735	0.841269841
Behaviour	Client018_j	peaceful	36	0.303336704	0.771929825
Personality	Client018_k	peaceful	42	0.293458619	0.762295082

Table 6.2: Sample results of emotion extraction and writing style features for client's sessions

6.3.2 Simulating "Classroom Behaviour" from the Dataset

One of the purposes of collecting a new dataset was to be able to test the feasibility of measuring the behaviour of a student when compared to the rest of the classroom, and of the classroom as a whole: this aspect is our research is particular important, as there are not very many studies focussing on social learning analytics [60, 78] and looking at the emotional factors, but this is of course an important element missing from the Motivational Interviewing corpus.

To this aim, we produced a simulation from our dataset, by grouping and labelling a series of sessions. We start from the assumption that in an online course we essentially have a number of students, a cohort, that is going through a specific topic together, and that course will be divided in well balanced periods (weeks, terms, etc). While each student has their own journey, and do not necessarily interact directly with the other students, unless there is any groupwork, the conversations in the classroom and with the teacher around the topic, will follow the same pattern. From our dataset, we obviously have that each session and each dialogue is very specific to the needs of the single client, however, given the nature of the motivational interview approach, each conversation will explore the issue further and build on the previous one, just as when learning a topic.

Given this assumption, we want to see if we can group together sessions that have been labelled with the same topic in the corpus for a client, gather together a number of clients who were exploring the same topic with their therapist, and "align" them together, as if they were all part of the same group. We disregard the content of this conversation, as of course different clients will have been talking about different topics, and we only consider the vector of values identified in the previous section for each session, while also grouping together sessions to form a series of "terms" or periods of time. Finally, we provide a vector of value to the whole cohort, but calculating the same metrics but applied to the collation of all clients' session together, and use this as a measure of how the textual output of this "cohort" as a whole.

The main topics identified on the 76 records of clients we extracted are *"Behaviour, Development, Relationships, Culture, Ability, Personality, Health"*, and an example of what we obtain, for instance, for the topic *Relationship*, by collecting together 20 clients with the same number of sessions, considering 4 "periods", and maintaining the temporal order between each session for each client, is a set of vectors, calculated as shown before, like the ones shown in Tables [6.3],[6.4],[6.5],[6.6].

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Relationships - Session 1	Client004_a	sad	40	0.33333333	0.49593496
	Client011_a	friendly	64	0.29015919	0.83333333
	Client016_a	peaceful	41	0.31590258	0.66942149
	Client019_a	sad	42	0.3358271	0.48031496
	Client031_a	friendly	26	0.31234867	0.61417323
	Client032_a	sad	9	0.2847769	0.28037383
	Client034_a	friendly	32	0.39128035	0.63636364
	Client110_a	friendly	19	0.39380197	0.58947368
	Client112_a	peaceful	22	0.26323676	0.51485149
	Client123_a	admire	14	0.40567376	0.50980392
	Client124_a	peaceful	20	0.42585551	0.68055556
	Client137_a	alert	30	0.36671725	0.53846154
	Client219_a	peaceful	38	0.27243844	0.76428571
	Client222_a	alert	8	0.35379464	0.5308642
	Client224_a	sad	19	0.43661972	0.40131579
	Client405_a	sad	45	0.31047266	0.48201439
	Client417_a	friendly	32	0.30936073	0.42222222
	Client419_a	peaceful	17	0.4200542	0.5915493
	Client420_a	peaceful	20	0.31132075	0.64210526
	Client421_a	friendly	23	0.37980769	0.60792952
	COHORT	**peaceful**	**95**	**0.251930262**	**0.543103448**

Table 6.3: Clients at their first session under Topic Relationships

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Relationships - Session 2	Client004_b	sad	33	0.31067961	0.50641026
	Client011_b	friendly	126	0.26078799	0.88461538
	Client016_b	friendly	14	0.43384982	0.62711864
	Client019_b	peaceful	4	0.40924464	0.29090909
	Client031_b	admire	43	0.33153275	0.53125
	Client032_b	sad	18	0.24479541	0.32338308
	Client034_b	peaceful	25	0.3545611	0.75609756
	Client110_b	peaceful	39	0.40437158	0.71764706
	Client112_b	alarm	13	0.31298905	0.51639344
	Client123_b	peaceful	15	0.36507937	0.59868421
	Client124_b	alert	40	0.29547229	0.62096774
	Client137_b	friendly	40	0.3564317	0.44262295
	Client219_b	peaceful	15	0.41621622	0.59459459
	Client222_b	peaceful	15	0.41421144	0.60526316
	Client224_b	peaceful	13	0.43168605	0.44642857
	Client405_b	sad	37	0.3315	0.59821429
	Client417_b	friendly	13	0.32740214	0.37719298
	Client419_b	peaceful	9	0.43933589	0.54794521
	Client420_b	sad	64	0.31663788	0.73684211
	Client421_b	peaceful	22	0.3998968	0.58762887
	COHORT	**sad**	**124**	**0.260240964**	**0.549050633**

Table 6.4: Clients at their second session under Topic Relationships

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Relationships - Session 3	Client004_c	sad	22	0.4260355	0.41509434
	Client011_c	friendly	143	0.25639518	0.75294118
	Client016_c	friendly	46	0.34367869	0.64571429
	Client019_c	sad	6	0.29503106	0.2970297
	Client031_c	peaceful	20	0.37112011	0.5974026
	Client032_c	alarm	15	0.29344074	0.3984375
	Client034_c	peaceful	25	0.36147757	0.75308642
	Client110_c	alert	37	0.30006835	0.46564885
	Client112_c	sad	23	0.35880708	0.42268041
	Client123_c	sad	10	0.32725766	0.59504132
	Client124_c	sad	38	0.33865506	0.728
	Client137_c	alert	22	0.47411616	0.5505618
	Client219_c	sad	23	0.3315508	0.48888889
	Client222_c	peaceful	8	0.31368697	0.45045045
	Client224_c	sad	34	0.37070778	0.45535714
	Client405_c	peaceful	37	0.31032216	0.50980392
	Client417_c	peaceful	12	0.30009775	0.36363636
	Client419_c	peaceful	14	0.45905707	0.47435897
	Client420_c	friendly	10	0.33677522	0.60869565
	Client421_c	peaceful	22	0.34336963	0.53246753
	COHORT	**peaceful**	**61**	**0.26454616**	**0.464516129**

Table 6.5: Clients at their third session under Topic Relationships

Topic	Session_name	Emotion	Emo_value	Richness	POS
Relationships - Session 4	Client004_d	sad	9	0.38132296	0.39759036
	Client011_d	friendly	142	0.26552655	0.82051282
	Client016_d	peaceful	49	0.38288719	0.72535211
	Client019_d	sad	15	0.37931034	0.36206897
	Client031_d	friendly	27	0.3345021	0.46341463
	Client032_d	peaceful	22	0.25981308	0.3697479
	Client034_d	peaceful	35	0.3661032	0.62441315
	Client110_d	peaceful	34	0.33647059	0.50458716
	Client112_d	sad	18	0.29724277	0.51485149
	Client123_d	indifferent	16	0.33661202	0.4950495
	Client124_d	alert	11	0.35609244	0.61363636
	Client137_d	sad	25	0.43613933	0.58064516
	Client219_d	peaceful	50	0.27782324	0.59803922
	Client222_d	admire	12	0.32696391	0.45679012
	Client224_d	peaceful	30	0.34348047	0.47204969
	Client405_d	friendly	31	0.36331939	0.39473684
	Client417_d	friendly	19	0.25109745	0.42739726
	Client419_d	alert	12	0.45131846	0.56140351
	Client420_d	sad	29	0.34527873	0.62765957
	Client421_d	peaceful	21	0.34468381	0.59259259
	COHORT	**peaceful**	**70**	**0.237706093**	**0.477794793**

Table 6.6: Clients at their fourth session under Topic Relationships

Now that we have completed the dataset, we can see what tasks can be performed with it, as described in the next Section.

6.4 Experiment 3: the Emotional Dashboard

Once all machinery is in place, this gives us the opportunity to look at a set of client sessions, and to use them to simulate the behaviour of a classroom. As explained above, the working hypothesis is that each "topic" of discussion is analogous to a topic in the classroom, and that each of the clients will react to different topics in a different way. The observer system we hypothesise in our conceptual framework will be able to perform an emotion/sentiment analysis and a stylometric analysis, and decide if there is a situation to flag, as shown in Figure [6.1]. In this session we present our informal experimentations with a set of different strategies for dashboards and visualisation techniques, which we have also presented at two educational conferences (ITiCSE[1] and EDEN[2]) [6, 7] for feedback from practitioners.

The first obvious task to perform is to explore the data for individuals, and track their changes overtime, to understand whether a change happened, and if this can be associated to a specific topic (Figure [6.5]). Ideally, we can look across the same session/topic and between different sessions/topics, and create a baseline for the same student which could help identify substantial changes in style or emotional behaviour, which the Observer can flag to teaching or pastoral staff.

Another task might be for the educator to understand if some topics are more prone to arouse negative (or positive) emotions in the class, therefore one might want to explore what are the main emotions that have been observed for each topic (Figure [6.6]).

There might be many variations of the same type of tasks, but the key notion that we want to explore in this Chapter is the notion of "change". We mentioned that one of the objectives we wanted to focus on, by identifying writing style features from conversations, was to be able to spot students who were "outliers" in terms of the general behaviour of the

[1]ACM Innovation and Technology in Computer Science Education (ITiCSE) `https://iticse.acm.org/`.
[2]European Distance and E-Learning Network `https://www.eden-online.org/`.

Sum of Writing Features	Column Labels				
	Ability	Behaviour			Relationships
Row Labels	Sad	Indifferent	Peace	Sad	Sad
client004_a	0.746047534				
client004_b					0.829268293
client004_c					0.817089868
client004_d					0.841129843
client004_e					0.778913319
client004_f				0.741380684	
client004_g			0.832413057		
client004_h		0.853142989			
client004_i	0.650902512				

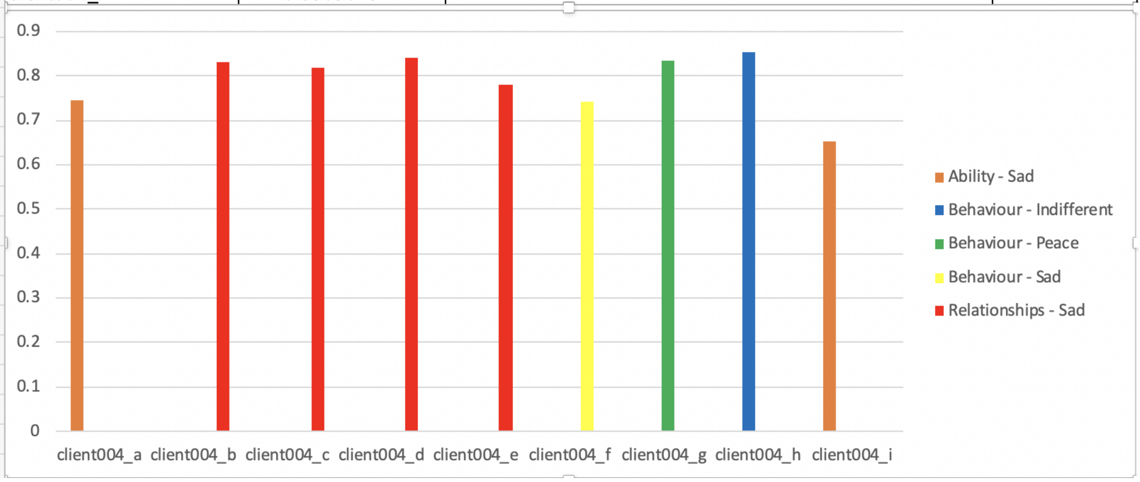


Figure 6.5: Example of a Dashboard to study the association topic/emotion for an individual (Client004)

class. Change could be defined in many ways, and in order to test a workable measure, we introduce a simple procedure in Algorithm 3.

Row Labels	Max of Postagging	Max. of Richness
Ability	**0.90797546**	**0.614583333**
Admire	**0.492753623**	**0.478079332**
Client137_d	0.492753623	0.383789587
client406_o	0.483870968	0.478079332
client425_f	0.275449102	0.315300546
Alert	**0.642384106**	**0.487323944**
Client124_p	0.642384106	0.3125
Client228_c	0.444444444	0.287679426
client419_n	0.452830189	0.487323944
Brave	**0.666666667**	**0.545112782**
client406_a	0.666666667	0.545112782
Cheerful	**0.618181818**	**0.451476793**
Client112_h	0.520833333	0.346740638
client419_c	0.618181818	0.451476793
Friend	**0.83203125**	**0.41560219**
Client011_a	0.824940048	0.112697198
Client011_b	0.83203125	0.229634672
Client031_b	0.552238806	0.281586022
Client031_e	0.601851852	0.379844961
Client031_m	0.436781609	0.29956427
Client137_a	0.603305785	0.41560219
client418_c	0.508928571	0.4137577
client425_c	0.282352941	0.334569733
client425_d	0.382488479	0.318044659
client426_a	0.296	0.24757953
client426_e	0.23255814	0.288987435
[illegible]	[illegible]	[illegible]

[illegible]	[illegible]	[illegible]
client426_d	0.330935252	0.284274194
Behaviour	**0.833333333**	**0.604**
Admire	**0.578431373**	**0.526636225**
Client032_f	0.302083333	0.322368421
Client205_c	0.578431373	0.329054054
Client221_d	0.489583333	0.384848485
client406_n	0.406779661	0.526636225
client419_f	0.31547619	0.38121118
Alert	**0.653846154**	**0.345953003**
Client124_i	0.653846154	0.345953003
Brave	**0.345323741**	**0.3372835**
Client220_b	0.345323741	0.3372835
Cheerful	**0.666666667**	**0.452418097**
client410_a	0.666666667	0.452418097
Friend	**0.733333333**	**0.466192171**
Client124_b	0.602272727	0.327403643
Client203_a	0.733333333	0.362644416
Client204_b	0.392307692	0.377850163
Client210_b	0.64556962	0.423390752
client420_p	0.68852459	0.343945069
client426_b	0.262357414	0.299270073
client426_c	0.571428571	0.466192171
client426_h	0.321428571	0.24160632
Furious	**0.613861386**	**0.283070596**
Client201_a	0.613861386	0.257537688
Client205_h	0.45	0.283070596
Indifferent	**0.497409326**	**0.355733662**
client004_h	0.497409326	0.355733662
Not Available	**0.4**	**0.555555556**

Figure 6.6: Example of a Dashboard for the association topic/emotion

Algorithm 3 Determining if a Change occurred

if $C.emotionType \neq Avg.GEmotionType$ **and** $|C.Richness - Avg.GRichness| \geq x$ **or** $|C.POS - Avg.GPOS| \geq x$ **then**

  change= True

else

  change= False

end if

Where:

C = Client,

x = threshold value,

$GEmotionType$ = General Emotion Type of the class,

$GRichness$ = General value of Richness,

$GPOS$ = General value of Part of Speech tagging.

To demonstrate how this might work, let us consider as an example from the group of sessions in our dataset, with the first session in Table [6.3] and following. From the data we can see that there is a group of clients who differ from the positive general emotion of **(peaceful)** in the session. If we apply Algorithm 3 and set a threshold $x = 0.05\%$ we are able to create a map of changes, across sessions, and for all students, in order to flag students who are drifting away from the general class behaviour, like the one in Figure [6.7], reporting clients for whom a change can be identified (cells labelled "TRUE" when a change is detected, and also highlighted in yellow when a change in all three components is detected) which shows, for instance, that client124 and client 137 behave differently both from the group on the same session and from themselves on their first session, so could be candidates for alerting to the teacher.

	First Session			Second Session			Third Session			Fourth Session		
	Emotion Compared with GEmotion	Richness Compared with Grichness	POS Compared with GPOS	Emotion Compared with GEmotion	Richness Compared with Grichness	POS Compared with GPOS	Emotion Compared with GEmotion	Richness Compared with Grichness	POS Compared with GPOS	Emotion Compared with GEmotion	Richness Compared with Grichness	POS Compared with GPOS
client004	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE
Client011	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE
Client016	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE
Client019	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE
Client031	FALSE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE	FALSE
Client032	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE
Client034	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE
Client110	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE
Client112	FALSE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE
Client123	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE
Client124	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
Client137	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
Client219	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE
Client222	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE
Client224	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE
Client405	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE
client417	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE
client419	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE
client420	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
client421	FALSE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	TRUE	TRUE

Figure 6.7: Example of a Dashboard flagging clients who are outliers

6.5 Conclusions

This Chapter concluded our set of studies to demonstrate the feasibility of the framework presented in this thesis, by describing how to attempt putting together a mechanism for creating an emotional baseline for a student, which would enable the emotional observer to notice if something has changed during the studies, and also to measure the status of a class as a whole, and in terms of students who could be outliers.

The technical limitation of the experiment, as with the others, is that the algorithms used were not optimised to the task; for example, in the first experiment: the emotion detection approach we found in off-the-shelf tool was not entirely fit for purpose, it was mostly based on a keyword-spotting approach which performs less with complex sentences, with a complicated structure and figurative language, while the main limitation in the second experiment are coming from the data itself as it does not show as much interaction among participants in the module.

Furthermore in the third experiment: if the system was going to be implemented for real then effort should be spent on deciding the most appropriate stylometric measures to be applied to classroom discussions, as well as creating a more fitting set of emotion keywords. In the original study by Whissell [216], improved further in [215], a Dictionary of Affect was created by a study with human participants who were asked to score words with their emotional load coming from literary corpus. This exercise would need to be repeated with a more relevant corpus, in order to provide more meaningful results.

The main limitation of the third experiment is of course that the dataset was a simulation of a classroom, so the aggregation in cohorts is only fabricated, as well as the identification of topics. Also, by coming from counselling session, the corpus is very emotionally loaded: again, a more relevant corpus should be collected directly from the classroom, but this poses many ethical challenges [178].

As a way to address this issues, and to evaluate the level of acceptability of this proposal with practitioners in the field, we organised a focus group studies, which will be described in the next Chapter.

Chapter 7

Focus Group Evaluation of the Approach

7.1 Introduction: Study Overview

One of the crucial issue of implementing a system like the Emotional Observer we have conceptualised, is understanding how useful it might actually be, and this needs to be mediated by ethical issues around using data for observing students. It could be that students would resent this, as they think their privacy is violated. It could be that teacher would resent this, as they believe they would lose the human touch in their communications with students.

Uptake is an important issue that needs to be addressed: too often computational systems only focus on the technical aspects, and do not explore the environmental ones. It was therefore felt important to conclude our set of experiments by exploring this precise issue, and refer back to practitioners to understand what they feel. In order to this, we turn to qualitative research, and we decided to use the mechanism of Focus Group Discussions.

This Chapter will introduce briefly this mechanism, just as any other tools we have used in this thesis, then explain how the study was conducted, and discuss the results of the study.

7.2 The Tool: Focus Group Discussion (FGDs)

The focus group approach has been widely used over time for a wide range of disciplines, such as scientific, educational research, sociology, communication and media studies [41, 63, 116, 136]. It is defined as a qualitative and observational process to gather information on specific topic from a small group of participants with similar characteristics, rather than a representative sample of a larger population [41, 135]. This method is being used to discuss a particular topic aiming to enhance, change or develop services, product, brand or campaign [145].

The focus group method is viewed as a cost-efficient technique in comparison with other techniques [136]. The significance of a focus group is its ability to produce data based on group interaction cooperation. Therefore, the group participants should feel completely comfortable with each other and feel free to participate in the discussion [163]. Sometimes FGDs are seen as synonymous with other types of interviews, in particular the semi-structured interview, but usually such interviews occur with individuals [151].

According to [145], in the literature there are seven types of focus group discussion:

1. **Single focus group**

   It is most frequently used in various fields by researchers and practitioners [136]. A single focus group is where a single moderator is asking questions and all the participants as one group discuss the topic in one place.

2. **Two-way focus group**

   Two focus groups (each with its own moderator) collaborate together. One group discuss a topic and the second group observe the discussion of the first.

3. **Dual moderator focus group**

   With this format, unlike the first type, we have two moderators collaborating within the same focus group. It helps to overcome the distraction of a single moderator.

4. **Duelling moderator focus group**

   Also with two moderators, this form of focus group places one moderator against another in order to discuss opposite sides of the issue to make sure that all topics are covered.

5. **Respondent moderator focus group**

   This format involves one or more of the focus group respondents taking up a temporary role as moderator to increase the chances of asking questions and to get more responses.

6. **Mini focus group**

   This features a small group of respondents, usually four or five. A smaller group of people provides a more friendly conversational atmosphere in contrast to other types of focus groups.

7. **Online focus groups** Sometimes called "Remote" focus group. It is not a different type but it used as adaptation of the traditional face-to-face methods, utilising the Internet to connect remote group members. It is done using phone or video conferencing, chat rooms or other online programs. This method makes it a good option to gather participants to discuss topics from multiple locations.

There are many advantages of using a focus group interview. First of all, it allows one to explore and test your hypotheses about a new product or service to solve any unexpected problems. It allows for exchanging ideas among participants as it is flexible structured discussion. Using focus groups could save time as it is the quickest way to get beneficial information and observe key messages that will help to take a decision about your product. In addition, one could better understand negative and positive opinions about your new product, design idea, or services. It is a way to gather trusted data from a potential expert.

It is important to realise that the focus group approach has some downsides. It provides qualitative data and cannot be used for quantitative purposes such as testing for a large population survey. Also, it is not recommended to use it for sensitive issues, as respondents will hesitate to talk freely and express their feelings and experiences. The data generated from the discussion could sometimes be difficult to analyse.

The focus group experiment conducted in this thesis represents, in the lifecycle of a potential system implementing our approach, the stage after the concept was identified and explored theoretically, but before committing to an implementation, and proceed to a complete system analysis and design. The concept was presented to experts and practitioners for feedback on its potentials, and comments on possible uptake. Therefore, we are not at a stage where we need to involve potential users for a complete requirement analysis exercise, before starting the deployment phase and make sure the system is fit for purpose, but rather at the stage where we need to be satisfied that the concept itself has potentials, by collecting opinions on the concept, after it was formulated in a framework, and studies were conducted to reassure the participants of its feasibility. This is also why we opted for a mini focus group approach, by gathering opinions from a small set of invited experts in online education. And, as participants were from different countries, we opted for an online approach, by using an e-conference tool.

7.3 Process of FGDs and Essential Steps

Focus groups are effectively a conversation with your potential participants to investigate a concept. The process consists of four main steps [137].

In the first step, the process begins by defining the research goals, purpose, and planning a moderator guide or discussion guide. We must consider what information we need to gather, and then decide how to do it. The questions should be appropriate and match the research goals. Then, we should consider the group participants; Should they have similar backgrounds or experiences? The participants should be collected from a community that you think will give you the best data.

The second step is the data collection, where the participants are grouped together over a period of time. Usually the time for the group to meet should not exceed two hours. Members in a focus group session are often encouraged to take their time to consider the subject that is under discussion. The place of the meeting should be comfortable and private where all participants feel free to talk. The moderator leads the session conversation by making sure that all the participants have arrived, and then starting the introduction and discussing the content. Next we ensure that all participants take part in the discussion by asking appropriate questions. The moderator tries to keep the conversation on track so that the groups do not lose their interest and are helped to maintain a focus on the topic.

The third step is to analyse the focus group data. The conversation session might be recorded, or a transcript may have been taken by a note-taker, or the moderator may summarise the findings. They would then write any additional notes or any observations that happened during the session.

Finally, we have the results and reporting step where the results and findings are published. All the data should be analysed in a coherent report in order to present the findings in an organised fashion. The interpretation of focus group data could take a wide variety of forms. The findings of the focus-group interviews can be described in an uncomplicated manner using relevant participant quotations [163]. This helps to make sense of the individual quotes and see how they relate to the findings, but also to identify connections between the data as a whole.

7.4 A Focus Group Meeting to Evaluate the Conceptual Framework

We describe here how we have implemented the four steps of the FGD, as depicted in Figure [7.1].

1. **Determine topic and goals:**

   We want to establish (1) how mental health is perceived as an issue by the practitioners (2) to what extent they use learning analytics or other automated tools to understand whether a student or a colleague is struggling (3) what is their reaction to the proposed framework, especially with respect to barriers to its uptake.

2. **Identify questions:**

   We would like participant to provide some feedback on the theoretical framework, by drawing on their practice and expertise, so we identified some prompting questions and follow up. The list of the questions asked in the focus group are represented in the "Interview Schedule" sheet, see appendix [G].

3. **Identify participants:**

   The invitation was sent to three professionals in the field of e-learning, two long term educators, who are also either Director of Studies or Lead of the Faculty in online programmes, and one who has lead the Student Experience department of online programmes, and who has therefore an oversight of both the student experiences and the type of questions online instructors are facing on a daily basis. We started by asking each of them to introduce themselves, briefly describe their job/area of expertise, and what their contribution might be to the discussion, appendix [H].

4. **Prepare location and select time:**

   A focus group was held on Monday, April 20, 2020, from 4:30-5:30 p.m. It was held online, conducted via Microsoft Teams, which allows communication and collaboration through groups, or teams `https://products.office.com/en-us/microsoft-teams/group-chat-software`. The session was recorded to aid the production of transcripts.

   The participants agreed to take part in the study by signing the "Participant Consent Form" which is attached in appendix [E]. The transcripts were anonymised. The video was not made public but was destroyed, subsequent to a transcript being made of the conversations. All the information about the study was made available on the "Participant Information Sheet", Appendix [F]

The study underwent full ethical approval according to the University of Liverpool regulations, See Appendix [D].

The full transcripts of the dialogue are in Appendix [H], and will be summarised in the following section.

7.5 Outcome of the Focus Group: Discussion

According to our brief, the conversation covers three main topics: the issue of mental health, the appropriateness of the use of Learning Analytics, the feasibility of the framework.

Participants described themselves as follows:

- **Participant A:**

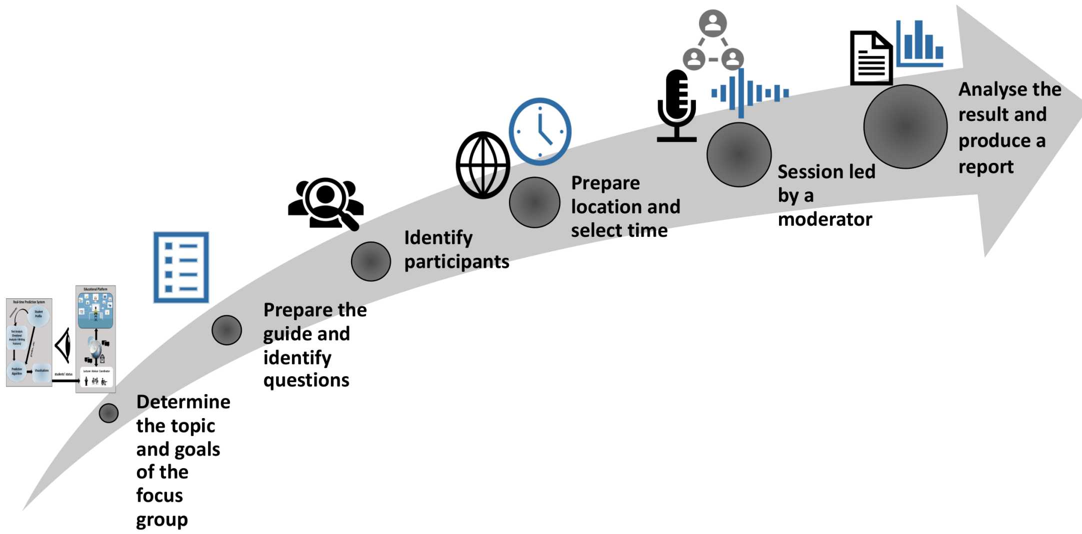


Figure 7.1: The steps of the focus group discussion technique

"I have been involved in online teaching and learning since 1995. I started working in the commercial area in online format and then eventually moved to the academic world in 2002. I am currently working as a lecturer in Higher Education and innovative use of technology in education. I have been involved in the overseeing [University name omitted] online programs, I worked on master's degree program and computing, so I looked at online teaching and learning from two different disciplines perspective. I also worked in online learning for my own campus environment as well as working in one of the other partnerships, the [University name omitted]. Then, in another commercial, for profit institution in the state of [name omitted] as well. I had a kind of interesting background in terms of working for different types of institutions. I have been working in the commercial IT area, first for

Microsoft and then continued to do research, conduct research in the notion of online presence for faculty and students."

- **Participant B:**

  "I had a number of years working on campus with students, with some teaching, but mainly in a support role. Then, I worked for 14 years in online master student primarily in student support and leading teams of advisor who are available for students studying online. I could offer some insight to how best able to do that how they support the students and what he saw on the campus worlds. Also, I have spent some time looking at predictive analytics and that's an area of increased research."

- **Participant C:**

  "I worked for a university face to face and retired in 2013, and during that time I was doing some part time online work for the [University name omitted] until 2013. Then, I became full time as the program director for the director of online studies there. And in 2019, I moved to [University name omitted] as a full time program director of their master's programs and their doctoral program. So, I have had some significant experience on ground teaching and about 20 years worth of online program."

7.5.1 On Mental Health

All the participants describe their experiences of the student mental health state. It has become more acceptable for students to come and talk about their concerns or their mental health issues. The respondents noticed there is an increase in the volume of students who come and talk about suffering from mental health issues.

Participant A for example said:

> *"I definitely have had students who have come to me with mental health issues; they talk about the stress of trying to manage their family or trying to manage their work. I have had students actually come to me and say, I have to take a break. I have a mental health document that I will share with you, but I ask for privacy, things like that."*

Participant B said:

> *"I would agree with Participant A. There was certainly a lot of cases or a number of cases where students would come forward and talk quite candidly about their experience; what they were going through."*

Participant C raised a concern about what caused the mental health issues. Students' mental issues might be different and affected by the context or a specific event. Particular contexts that create stressful and increased levels of mental health concerns are shared by many students. As mentioned by **Participant A** for example, a night before the submission date of an assessed piece of work. Also, **Participant B** mentioned that technology barriers would raise mental health issues. Some students get nervous or anxious, for example, when the server went down; especially if there is a submission due.

Participant C said:

> *"If the student cannot meet deadlines because of health issues, they have to go and do a petition. They have to do an exception and then have to wait for a few days, maybe a month or two to really find out what has happened to them. A lot of the mental health, I think is exacerbated because they can not find out what is going to happen to them right away."*

Immediate interventions regarding student mental health or health issues help the student as **Participant C** said, when referring to her personal experience "immediate petition's experiences" example when she compared the rule taken by two Universities they worked with when they deals with student health issues.

Participant C argued that the online environment is less stressful than the traditional face to face class. But as **Participant A and Participant B** mentioned, it is clearly indicated that online factors tend to increase worries or stress among online stakeholders.

Participant A said that:

> *"I am wondering, one of the things that we are talking about is that particular online context, in which we are working somewhere where it may be creating stressful and increased levels of feeling of mental health concerns."*

Participant C said:

> *"I think a lot of the times, the experience I had was that students would talk about some of the perhaps mental health issues that they face. They would come up when those kind of additional factors happened (the library service was down or the assignment is not clear) and you hear and see the anxiety, the outpourings, long emails or distressed phone calls."*

Teachers or student support might disclose some particular issues that students face even if they do not declare something. For example, **Participant A** said that *"I have picked up something, a mess that is not typical as a student and often it is near the end of the module."*. Also **Participant C** mentioned that *"You do get the feeling that a student is not in a good place, you can see behind what they have written; that they are in a very bad place and they really need to step back and talk to somebody about it."*

Participant B raised the point that:

> *"We get to know students more through their online behaviour and their online work because everything is visible."*

7.5.2 On Learning Analytics

When the participants talk about the online environment, it becomes clear that their actual experience reveals that **online messages provide a clear way to see what is going on from the writing**.

Participant B mentioned that:

> *"I think one of the things that is absolutely key is of course, it is sentence because it is recorded. So I had a case two years ago, where a student wrote to us to tell us that he was in a very, very bad way, he was going to commit harm to himself."*

They also talk about how important it is to build that relationship between the student and their supporter. What is vital is, as much as possible, to have students supported by the same person throughout their journey. Being able to recognise any changes in the student's behaviour produced a much more personalised intervention with a student while they are in the middle of learning. There seems to be a belief among participants that **knowing the individual students very well will help to identify anything that might be seen as a change in behaviour or a sign of concern.**

Participant C said that:

> *"It kind of stands out, if you have been doing it for a while and you have been tuned in to the student. You know when something is different.* ***Participant A*** *had the same approach to detecting changes in a student "Often it is how they are write things. Their style of writing. How they phrase, how they present an*

> *idea. If their typical approach is to be logical and then they stop writing it in logical format that you are used to seeing."*

Participant B discussed the role of using the predictive analytics to efficiently identify what the particular behaviours. He said that

> *"I think the idea that we can potentially know the students online through all of their footprints "the millions and millions of footprints" left by every single click."*

Participant A supported the benefit of using LA by saying:

> I have understood right now looking at learning analytics as a vehicle for real time interventions within the learning platform. To produce a much more personalised intervention to a student while they are in the middle of learning as opposed to understanding the student on an ongoing basis; having a profile at the start and then predicting how they might perform down the road.

Not every instructor wants to know how learning analytics measures student data. **It is difficult to monitor every student's learning pathway, especially if we have large number of students** as **Participant C** said *"We have a lot of students. So the only thing the instructor sees is what that student brings to that classroom today for this week, or eight week term."*

7.5.3 On the Ethical Aspects

Greater attention is placed by the respondents when they talk about the ethical issues related to our model. Ethically, this approach is potentially very challenging even though the intention of the research is to enhance online teaching and learning. What is difficult about it is what was mentioned by **Participant A** *"They are related to categorising or* *prejudging students."*. **Participant B** *"I imagine all of the potential pitfalls and dangers of living in a world where online personality can be relatively accurately, and maybe not entirely accurately predicted, as all."*

Participant B said:

> *"I think it is really the most important issue around. I do not think students signing up for a study at an online university are possibly even necessarily aware how much data you can gather and are able to gather from them to build up your model of who they are and what their behaviour might be. It is a super strong point, and ethically very challenging to get the balance right, I think between support and invasion of privacy. All kinds of privacy."*

If the model works for instructors monitoring students it will be easier to defend the privacy issues. Participant C said:

> *"This could be more predictive for a faculty member. And we do not have that many ethical issues."*

7.5.4 On the Conceptual Framework

The respondents were asked what they think about the proposed framework and the subsequent flowchart of the text analysis steps. The system, as mentioned by the participants, **would help anybody who is trying to keep track of how well faculty members or students are doing while they are working alone. Participant B** said:

> *"What I saw is really important to the support staff for building one to one relationships and getting to know the personalities, and getting to potentially prevent or see signals of future difficulties."*

Participant B supported the usefulness of the framework *"I do think that this will be useful prediction."*

If the system addresses the issues for faculty members or students it would be something that would be helpful. It seems that still **it is important to have human follow up even if those flags would be as accurate as possible.** Whatever outcome the algorithm produces we leave the decision to teachers to intervene. The proposed system would be useful as mentioned by **Participant A** *"If I take myself out of the current platform, and put myself into what if we had an educational platform that allowed real-time knowledge about students and shifts in students behaviour through the use of analytics. What potentially could that do in terms of improving our ability to intervene in a more timely basis, or potentially in a slightly different basis than what we would have ourselves have even thought about."*

Interesting information can be retrieved from discussion about the visualisation aspect of the dashboard. The participants are asked directly to speak from their experience as to what they think about the future output of an ideal system. **All participants extensively agreed that the system would benefit from presenting a single score, that would enable the e-learning mentor to decide who they needed to look at in more details.** There seems to be a belief that as soon as the system presents multiple scores, the ability to decide who you want to try to look at would be impossible. **Participant B** agreed that *"If you are focusing on this one issue, you want essentially a single score as much as possible for any particular student."*

7.6 Conclusions

We have highlighted in bold-font in the previous section the passages of the conversation that we believe provide more insight into the evaluation of our system. We summarise them here:

- Participants already feel a strong sense of the presence of mental health problems among students. The participants noticed that students' mental health is a growing concern and the number of cases they came across was obviously more than in the past. This could be because we have technology that allows students to talk about it in a way that maybe they did not do historically. They feel more connected now because they can meet beyond just the text.

- Participants generally understand that on-line learning is kind of a stressful and worrying environment. It could be more difficult for some students, and lead to emotional distress because of many factors. On-line learning requires self-monitoring and self-motivation and in consequence could have negative effects. Students' ability to control their behaviour, cognition, and motivation requires support from their universities. Consequently, the use of learning analytics is becoming important in higher education because it is opening up a new way to support students by analysing their learning behaviour.

- Participants identified the importance of predictive analytics or learning analytics and how we use the past data of students to support their learning. These data are gathered to monitor students' performance and help to implement powerful learning formats to optimise the learning experience. Indeed, there is an increase in using tracking systems in e-learning environments as a result of the willingness of researchers to enhance on-line teaching and learning experiences. Universities use statistical data for students as individuals or groups in order to support their learning outcomes. They are also confident that the way in which the students is informative about their mental state, and a change in their style a powerful indicator that something is going on.

- Participants recognised the ethical issues related to sensitive data automatically recorded in an online environment. And they identified how we consider data privacy regarding user tracking and personal data usage. Ethical considerations should include significant information describing how your data is used and who has access to it. One interesting indication raised in this discussion is related to categorising or prejudging students according to their history of learning activities and their profiles. This ethical concern is manifested in the potential wrongful discrimination of students as individuals of interest according to statistical risk. This could lead to that student being stereotyped and mistreated in the assessments. Some groups of students will be treated differently from other groups. This is indeed the case, but the concept here in our research is to use the tracking data and benefits of learning analytics to help keep students safe, and intervene only to identify a student at risk, not to exclude them from the programme of study.

- The tracking of students is a challenging task for the e-tutors. Therefore participants had strong levels of trust for a system that would help to provide an early intervention, and which could help to identify warning signs of student behavioural changes or problems, before these result in a crisis situation. They described their ideal system as a system that would show an alert flag for a student who behaves differently than usual. They addressed the importance of having one single score rather than multiple scores to focus on the issues of any particular student.

From the focus group, we are encouraged to have confidence in the fact that our approach is both appropriate, and our proposal very needed and timely. This concludes our series of experiments, and we are ready to draw general conclusions in the final Chapter.

Chapter 8

Conclusion and Further Work

8.1 Summary of Contributions

In this thesis, we studied how we could enhance certain aspects of the students' online learning experience, proposing a system aimed at incorporating an emotional "tracking" mechanism to follow students in a VLE. We discussed the pedagogical research underpinning this, and how experiencing positive emotions such as feeling secure, happy and excited, or on the contrary, negative emotion like frustration, fear and sadness, can enhance or disrupt the learning efforts of students.

We discussed how early identification of emotions/sentiment of students in an online learning environment help identify students that are about to struggle on the course or program of study, and that, while in a typical in-person class, it is easier to observe students and their reactions, in an e-learning environment observing these reactions is challenging. In addition, these type of environment, with the feeling of isolation that remote learning is associated with, many more mental health issues can arise.

In order to address this problem, a new system architecture was proposed, with the aim to observe and analyse student behaviour at real-time in an e-learning environment. The objective of the thesis was to study the feasibility of such system improving the efficacy of the e-learning environment, with the aid of well known AI technologies. A conceptual framework, together with a workable flowchart of how events are triggered in a classroom, was proposed.

The feasibility study gave a flavour of the techniques currently used to detect emotions/sentiment and other features from text, and while the selection of the tools was not intended to provide the most sophisticated results, we discussed the limitations of those tools, and recommendations for identifying suitable approached for emotions/sentiment extraction, as well as a proposal to combine social learning analytics and social emotion analysis to the interpretation and evaluation of classroom behaviour.

In particular, the findings from the first experimental study indicated that off-the-shelf tools for text emotion recognition could be utilised to extract the emotions from text, and that these could integrate with online learning platforms. The second experiment demonstrated how features could be selected from the analytics provided by online learning platforms in order to predict student performance or student final grade. The third experiment showed how to combine emotion detection and writing style analysis to detect changes in behaviour, both at individual and group level. Finally, as we noted the importance played by the level of acceptability towards the idea of using AI technology to help teachers observe student behavioural changes, an online focus group discussion with three experts demonstrated that practitioners would value a solution based on the analysis of student emotion as an aid to investigate any issues around student behaviour.

A summary of the main contributions of this work are therefore as follows:

- A novel architecture for an "emotional observer" system sitting on top of a general Virtual Learning Environment. This was fully evaluated from both the technical feasibility and the acceptability to stakeholders perspectives in what is, to our knowledge, the first complete feasibility study for a system of this kind.

- A roadmap for the implementation of such architecture, which draws on a combination of existing techniques and algorithms. The combination is new to the educational field, and provides a new way to look at student outcomes. The roadmap was tested on either available or especially built datasets.

- A new notion of "emotional tracking" of either an individual or a collection of individuals with respect to a task, able to provide longitudinal analysis of how emotions change overtime, and in combination with changes in the style of writing. Our approach using a combination of emotion, sentiment and stylometric analysis is novel, and showed great potential for identifying when something "unusual" happens for an individual or a group, which is especially important in e-learning for early detection of drop-outs or risk of failing the course.

- A new data structure and set of algorithms for providing such emotional tracking in a simulated classroom, incorporating features for emotion/sentiment analysis together with stylometric parameters, like richness of the text and complexity of the grammar. A dataset was created to flesh out this structure as an elaboration of a proprietary corpus from the field of Motivational Interviewing.

- A set of recommendations for the a visual "emotional dashboard", which could help faculty track students at risk.

8.2 Limitations of the Approach

In this thesis we took the ambitious view of going broad rather than specific, by addressing the overarching problem, rather than finding optimised solutions to specific tasks. Therefore, the limitations of the approach come from the fact that each issue would merit a study in its own right, going deeper into the analysis of the solutions.

The fundamental limitation of this study, as any other study in the field, is given by the lack of suitable datasets. The ethical implication of collecting extensive datasets, going beyond simple statistics on logging time and grades, and using them to categorise students, is possibly the hottest theme in educational research [178]. All our experiments would have been much more meaningful if we managed to access or collect a real dataset, from courses delivered to students progressing across different modules (so, not necessarily MOOCs) of whom a bigger picture could be derived by also including communication with personal tutors and so on. Of course, this was not possible, the ethical approval process to obtain this from any university we contacted was not feasible in the time span of this PhD.

For a more technical point of view, the emotion detection approach we could find in off-the-shelf tools was again not entirely fit for purpose, they were mostly based on a keyword-spotting approach, together with syntax based rules, but ambiguity in text related to emotion is a big issue, the classical example is the sentences "I laughed at him" and "He laughed at me" which carry very different emotions. We already discussed in the relevant chapters how more investigation would be needed to find the most suitable tool, possibly fed by a good corpus of student based conversations, which leads us back to the previous point.

Finally, while we discussed the technical feasibility, and the perceived acceptability of the framework, we could not address the issue of practicality of the approach, which would need to look at problem like integration with major e-learning platforms from computational perspective, or how learning and teaching agreements might need to change from a business/educational perspective, to insure and reassure students what level of privacy is preserved, and how giving it away will result in a measurable positive impact. This remains outside the scope of this work, although we recognise it is possibly the set of issues with highest priority to consider.

8.3 Directions for Future Research

The learning and teaching methods could be enhanced by analysing the learner data in certain ways. The central contribution of this research is the use of emotion features to detect student behavioural change in addition to learning analytics data. We present here a road-map for some possible future directions in the various topics we have touched in this thesis.

Further study is required to find methods, approaches, and techniques to study implicit emotional/sentimental statements. Modern machine learning techniques are available to get more accurate results in emotion detection, however ML is not always optimal for the training on complex sentences, which are the most problematic to classify with respect to emotions. Deep learning techniques (e.g. Convolutional Neural Networks(CNN)), have been used to recognise emotions from facial expressions or speech, it could be interesting to see how they perform in recognising emotions from text.

For sentences, while there is advanced research in sentiment analysis, fewer approaches concentrate on emotions and the combination of sentiment and emotions: we have addressed this in Experiment 3, but much more is needed. In particular, for complex sentences, for example those containing more than one emotion, identifying implicit emotional/sentimental requires a set of grammatical and semantics features. Most of the available techniques use an Hybrid Approach which combines lexicon and machine learning approaches. Further development of this would require better affective lexicons in conjunction with other techniques that extract effective features for sentiment and emotion classification, on the same line as resources for opinion mining SentiWordNet[1], SenticNet[2] and its evolution EmoSenticNet[3].

The combination of emotion/sentiment analysis and stylometric analysis is a strong candidate for robust further research, there are many stylometric metrics that could be investigated, also in combination with authorship detection, something what would appeal educational research on academic integrity and plagiarism. And the application of this to the cohort of students, rather than individuals, is also something requiring more research, especially in the context of Social Learning Analytics.

Prediction has been a strong drive for the use of learning analytics so far, with research concentrating on drop out or failing students, in the effort to improve attrition and performances, however we maintain that there are other features worth predicting, and the emotional and mental wellbeing of students one of these. The lack of large datasets for training remains an issue, but new models of machine learning so called "one-shot" or "few-shots", based on generalisation from few examples [210], are definitely something to explore further.

A last avenue for further research that we want to mention, but not the least by any means, is the one related to visualisation. The focus group discussion concluded with agreement that the online educator overload of information is an issue, and that any tool providing extra information related to the classroom needs to be very effective from a visual/attention perspective. In e-learning environments, information visualisation techniques supply useful information to Learning Dashboards, and more research is needed to find the most effective visualisation of an Emotional Dashboard like the one we envisage. When presenting our

[1] http://sentiwordnet.isti.cnr.it/
[2] https://sentic.net/
[3] https://www.gelbukh.com/emosenticnet/

framework to educators, we have considered various prototypes for instance for visualising the performances of a classroom and outliers, either in positive (green stars) or negative (red triangles), e.g. see Figure 8.1 which we presented in [6].

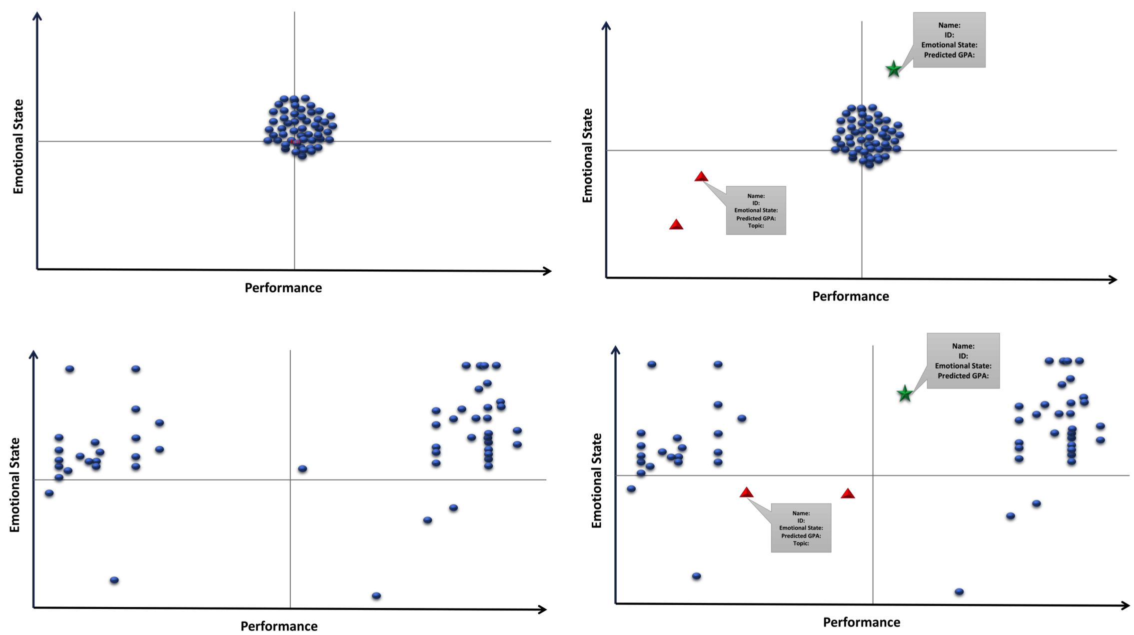
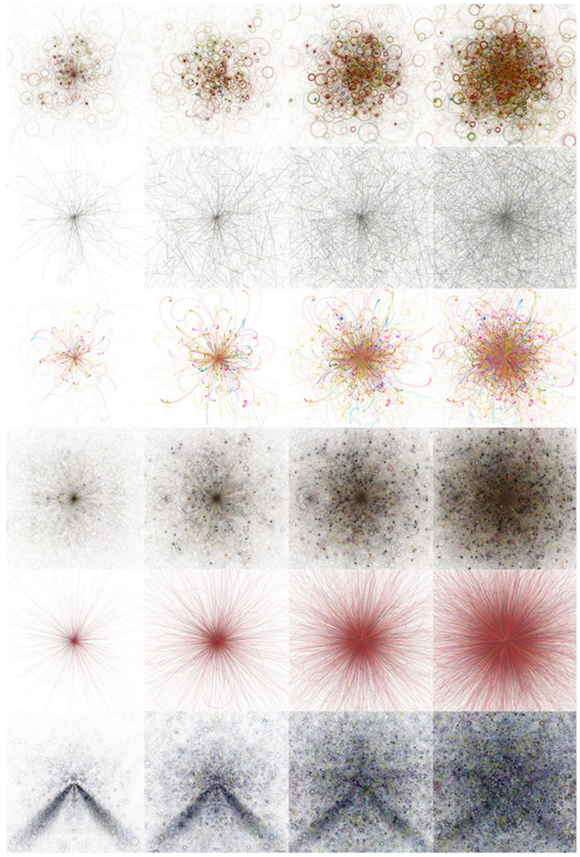
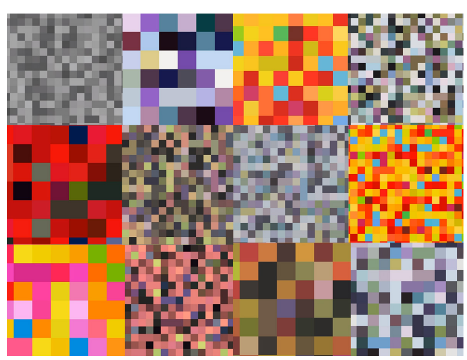


Figure 8.1: Visualisation prototypes to represent base line class data and outliers

In addition to classic analytics representations, it is worth investigating further the more creative visualisations of the emotional state, and its evolution. One good example, to stay in line with our choice of tools for the experiments, would be to see how Synesketch visualisations of emotion growth or intensity [105] (see Fig. 8.2) could be integrated with the learning analytics visualisations.



Growth of the six emotions (from top: Happiness, Fear, Surprise, Disgust, Anger, Sadness).



Intensity of emotions (from left to right, top to bottom: neutral / strong sadness / middle happiness / weak fear / strong anger / weak disgust / weak sadness / weak happiness / strong surprise / weak anger / strong disgust / middle sadness).

Both images are reproduced from `https://krcadinac.com/work/projects/synesketch`

Figure 8.2: Synesketch visuals of emotions growth and intensity

8.4 Final Remarks

UNESCO's *Beijing Consensus on Artificial Intelligence (AI) and Education* [204], the "first ever document to offer guidance and recommendations on how best to harness AI technologies for achieving the Education 2030 Agenda", was published in 2019 and adopted by representatives of member states as well as UN agencies, academic institutions, and other stakeholders in the private and public sector. Among the recommendations included in the document, we particularly note one:

> *Ensure AI technologies are used to empower teachers rather than replace them, and develop appropriate capacity-building programmes for teachers to work alongside AI systems.*

Despite the start of our work on this PhD predates this document, this is one of the driving motivations for this research: we believe in AI learning technologies that support and improve learning outcomes, that open up possibilities for real time feedback, but also that are designed for supporting and completing the role of the teachers, never to replace them. We hope that our work is a small step towards this important objective.

Appendices

Appendix A

Complete Results for Experiment 1

A.1 Complete Sample of the Records from ISEAR Dataset

.

Text from ISEAR dataset	labelled Emotion
I was selected to come here (University, College) when I was least expecting it.	Joy
When my brothers had passed all of their exams and were able to graduate from their courses.	Joy
When I was told that I was selected to attend Medical Assistant Training.	Joy
When I was accepted for my third year (G-10) at my former secondary school.	Joy
I was emotionally happy when I was in love with the girl I had longed for.	Joy

Text from ISEAR dataset	**labelled Emotion**
A neighbour's girl had disappeared and many people were looking for her. Someone had gone to notify the police. Something had certainly happened to her.	Fear
When I was at home alone, I felt a super-natural force, dangerous for me and the people close to me.	Fear
My fear appeared in the form of jealousy. I was afraid that my girl-friend had fallen in love with another man, I was afraid to lose her.	Fear
When a big angry dog put its snout on my arm and had I made one movement it would have bitten me.	Fear
A classmate told me I must have bribed the class leader to let me go to your English lecture.	Anger
I felt anger against a colleague of mine during a rehearsal in acting. He hadn't learnt the text of an opera act in the course of several months and thus making difficulties for the rest of my colleagues.	Anger
I had studied for almost one week for my physics- examination. With difficulty, I passed the exam. I was angry about the teacher and also about myself because I had not remembered enough during the exam and because the time that I spent studying was wasted.	Anger
When my car froze, and I could not start it.	Sadness

Text from ISEAR dataset	**labelled Emotion**
Several years ago my mother died. She had been ill for a long time, but nevertheless her death came unexpectedly. I did not and I still do not want to believe that it is true.	Sadness
When I feel helpless after having tried to help someone without any result.	Sadness
When my grandfather died, I saw my grandmother crying against my aunt's shoulder (I had never seen my grandmother cry before).	Sadness
In 1977, my grandfather, to whom I had a very close relationship, died.	Sadness
When I was about to clean the draining board and saw it looked a underneath the sink (I live in a students hostel).	Disgust
I was in a train when a woman started talking loudly and attracting everybody's attention. The worst thing was that she was discussing something, about which she knew nothing, with another person.	Disgust
My room-mate was drunk, he vomited on the floor and fell face.	Disgust
Once I was not ready for a seminar and I was asked to leave.	Shame
As a little girl, I was sick in the middle of a school day and I vomited in the basin of the classroom.	Shame
I felt this when I was copying homework for one of my classes.	Shame

Text from ISEAR dataset	**labelled Emotion**
My schoolmates were teasing a pupil who was not able to defend himself very well; I should have taken his part.	Guilt
When I finished a love affair where I was responsible of the sad end.	Guilt
In an exam I answered the questions rather carelessly and afterwards I thought that the exam would have been better had I answered more carefully.	Guilt

Table A.1: Complete sample of the records from ISEAR dataset

A.2 Complete Sample Results by Synesketch Tool

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
When I passed the TOEFEL with very good marks.	Joy	1	1	1	0.2	0	0	0	0
Passing an exam I did not expect to pass.	Joy	0	1	0.4	0	0	0	0	0
Joy of giving birth, and of sharing that joy with my husband. Moments of complete happiness and feelings of so much love.	Joy	1	1	1	0	0	0	0	0
When abroad, while driving a car along a dark, winding road.	Fear	1	-1	0	1	0	1	0.1	0

Text from ISEAR dataset	labelled Emotion	General weight	Valence	Happiness	Sadness	Anger	Fear	Disgust	Surprise
When, as a child, I was nearly knocked down by a car.	Fear	0.2	-1	0	0.1	0.1	0.1	0	0
A classmate told me I must have bribed the class leader to let me go to your English lecture.	Anger	0.0	-1	0	0.6	0.4	0.6	0	0

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
I felt anger against a colleague of mine during a rehearsal in acting. He hadn't learnt the text of an opera act in the course of several months and thus making difficulties for the rest of my colleagues.	Anger	0.0	-1	0.1	0.9	1	0.5	0.5	0.1

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
I had studied for almost one week for my physics- examination. With difficulty, I passed the exam. I was angry about the teacher and also about myself because I had not remembered enough during the exam and because the time that I spent studying was wasted.	Anger	0.0	-1	0.2	0.2	1	0	0.3	0
The death of a close friend.	Sadness	1	-1	0	1	0	0	0	0

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
I was told by a good friend that we couldn't be friends any more because of his relationship with another girl.	Sadness	1	-1	0	1	0.1	0	0	0.1
When I was told that a good friend was seriously ill.	Sadness	1	-1	0.2	1	0	0	0	0.3
My room-mate was drunk, he vomited on the floor and fell face.	Disgust	0.2	-1	0	0	0.1	0	0.5	0.1

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
When I was about to clean the draining board and saw it looked underneath the sink (I live in a students hostel).	Disgust	0	0	0	0	0	0	0	0
Once I was not ready for a seminar and I was asked to leave.	Shame	0.3	-1	0	0.1	0.1	0.1	0.1	0
As a little girl, I was sick in the middle of a school day and I vomited in the basin of the classroom.	Shame	1	-1	0.3	1	0.2	1	1	0

Text from ISEAR dataset	labelled Emotion	General weight	Valence	Happiness	Sadness	Anger	Fear	Disgust	Surprise
I felt this when I was copying homework for one of my classes.	Shame	0.3	-1	0.1	0.2	0	0	0	0
I experienced long ago when I was sightseeing Bulgarians in a foreign language	Shame	0.1	-1	0	0.1	0	0.1	0	0
Failing an exam because I did not work hard enough.	Shame	0.1	1	0	0	0	0	0	0
My schoolmates were teasing a pupil who was not able to defend himself very well; I should have taken his part	Guilt	0.6	-1	0.2	0.2	0.1	0.1	0.1	0

Text from ISEAR dataset	**labelled Emotion**	**General weight**	**Valence**	**Happiness**	**Sadness**	**Anger**	**Fear**	**Disgust**	**Surprise**
When I finished a love affair where I was responsible of the sad end.	Guilt	1	-1	1	1	1	0.2	0.2	0
When I shoplifted a pair of earrings from Coles and my Mum caught them in my bag.	Guilt	0.1	0	0.1	0	0.1	0	0	0
My friend came to the concert for my sake as it was me who had organized it and the concert was unsuccessful.	Guilt	0.8	-1	0	0.4	0	0	0	0

Text from ISEAR dataset	labelled Emotion	General weight	Valence	Happiness	Sadness	Anger	Fear	Disgust	Surprise
In an exam I answered the questions rather carelessly and afterwards I thought that the exam would have been better had I answered more carefully.	Guilt	0.1	-1	0	0	0	0.1	0	0

Table A.2: Complete sample of the results by Synesketch tool

A.3 Complete Sample of Failure Cases of Results by Synesketch Tool

Sentence	**Failure Categories**	**General Weight**	**Valence**	**Anger**	**Disgust**	**Fear**	**Sad**	**Happy**	**Surprise**
When my friends did not ask me to go to a New Year's party with them.	problem with negation	0.1	1	0	0	0	0	0.1	0
When I read a theoretical book in English that I did not understand.	problem with negation	0	0	0	0	0	0	0	0
When one's studies seem hopelessly difficult and uninteresting.	keywords not recognized	0.6	0	0	0	0	0	0	0
When I was not accepted as a student in finance and accounting.	problem with negation	0	0	0	0	0	0	0	0

Sentence	Failure Categories	General Weight	Valence	Anger	Disgust	Fear	Sad	Happy	Surprise
When I had not understood anything after a lecture.	problem with negation	0	0	0	0	0	0	0	0
A case of unrequited love.	problem with negation	1	1	0.1	0	0	0	1	0
Thoughts revolve around failing the subject and the consequences it would have for the future.	keywords not recognized	0.1	1	0	0	0	0	0.1	0
Not succeeding in a cross-country skiing competition;	problem with negation	0	0	0	0	0	0	0	0

Sentence	Failure Categories	General Weight	Valence	Anger	Disgust	Fear	Sad	Happy	Surprise
When I did not get the salary increase that I had been expecting and understood how little one's work was appreciated.	problem with negation	0.1	1	0	0	0	0	0	0
I heard that a former superior of mine had died;	keywords not recognized	0.3	1	0	0	0	0	0.1	0
My grandmother died, and my mother called me one sunday morning in the Autumn.	keywords not recognized	0.2	-1	0.1	0.1	0	0.2	0.1	0

Sentence	Failure Categories	General Weight	Valence	Anger	Disgust	Fear	Sad	Happy	Surprise
My girlfriend gave me the mitten (left me).	failed with metaphoric expression	0	0	0	0	0	0	0	0
My grandfather died;	keywords not recognized	0	0	0	0	0	0	0	0
My father had a heart attack when I was not at home (I was still living with them).	keywords not recognized	0.2	1	0.1	0	0	0	0	0.1
When I was told that a good friend was seriously ill.	confused the keyword	1	-1	0	0.3	0.2	1	1	0

Sentence	Failure Categories	General Weight	Valence	Anger	Disgust	Fear	Sad	Happy	Surprise
As a ten-year-old I was at the funeral of my grand-father.	keywords not recognized	0	0	0	0	0	0	0	0
After an exam which I failed.	keywords not recognized	0	0	0	0	0	0	0	0
When I understood that my marriage was falling apart;	keywords not recognized	0	0	0	0	0	0	0	0
A person close to me told me that his positive regard depended on my conduct.	keywords not recognized	0.6	-1	0.3	0.2	0.3	0.3	0.3	0
When I heard that a good friend had committed suicide.	keywords not recognized	1	1	0	0	0	0	1	0

Sentence	Failure Categories	General Weight	Valence	Anger	Disgust	Fear	Sad	Happy	Surprise
When one of my cat died of a disease.	the word disease help with the word cat	0.3	-1	0	0.1	0.1	0.2	0.1	0
I made a long-distance call to people rather close to me and I thought about the sad incident that had happened to them in the near past;	only concered sad keyword	1	-1	0	0.2	0.2	1	0.2	0

Sentence	Failure Categories	General Weight	Valence	Anger	Disgust	Fear	Sad	Happy	Surprise
She was very close to me.	the word close and past tense showed sadness	0.2	-1	0	0	0	0.2	0.1	0
I did not quite succeed in breast feeding my baby.	failed with negation	0.1	1	0	0	0	0	0.1	0
At a lack of love of my father for my mum;	failed with negation	1	1	0.1	0	0	0	1	0

Sentence	Failure Categories	General Weight	Valence	Anger	Disgust	Fear	Sad	Happy	Surprise
I was told by a good friend that we couldn't be friends any more because of his relationship with another girl.	failed with negation	1	1	0.1	0	0	0.1	1	0.1
When the guy I was in love with told me that he had met someone else and that we would not meet again for a year.	aggregation of many emotions make it difficult to judge	1	-1	0.9	0.7	0.9	0.8	1	0

Sentence	**Failure Categories**	**General Weight**	**Valence**	**Anger**	**Disgust**	**Fear**	**Sad**	**Happy**	**Surprise**
When I feel helpless after having tried to help someone without any result.	many keywords lead to high prediction values for all emotions	0.9	-1	0.9	0.7	0.9	0.8	0.8	0
When I feel at peace with myself and also experience a close contact with people whom I regard greatly.	Failed handling keywords	0.6	-1	0.1	0.2	0.2	0.2	0.1	0.1

Table A.3: Complete sample of failure cases of results by Synesketch tool

Appendix B

Complete Results for Experiment 2

B.1 Complete Features Used in Phase 1 to Predict Final Grade

Percent_grade	SessionLength(sec)	NumEventsInSession	Module_type	FinalGrade
100	13283456	235264	problem	Distinction
100	11873125	709544	problem	Distinction
100	4873033	58916	problem	Distinction
100	10445416	150388	problem	Distinction
100	19252548	410400	course	Distinction
100	789064	17914	sequential	Distinction
100	8005238	73868	problem	Distinction
100	25570846	823669	sequential	Distinction
100	13763520	328276	chapter	Distinction
100	3142352	71392	problem	Distinction
100	21825496	272426	video	Distinction
100	8656479	122616	problem	Distinction
100	41845342	982366	video	Distinction
100	8184918	137529	problem	Distinction
100	9864960	241080	problem	Distinction

Percent_grade	SessionLength(sec)	NumEventsInSession	Module_type	FinalGrade
100	2629604	36076	chapter	Distinction
100	14656488	172788	problem	Distinction
100	6446040	131400	problem	Distinction
100	14910060	128172	sequential	Distinction
100	5259870	69120	problem	Distinction
100	11755891	200158	problem	Distinction
100	8659770	71736	problem	Distinction
100	9741589	125356	problem	Distinction
100	82127699	1478728	chapter	Distinction
100	4025738	131306	video	Distinction
100	3358830	22660	problem	Distinction
100	7829040	87720	problem	Distinction
100	2965976	26104	problem	Distinction
100	6958560	199360	sequential	Distinction
100	1093540	6059	sequential	Distinction
100	645405	4080	problem	Distinction
100	504680	5270	chapter	Distinction
100	9692280	402192	problem	Distinction
100	278380	7130	chapter	Distinction
100	2249933	36792	course	Distinction
100	741356	7250	problem	Distinction
100	340290	2793	problem	Distinction
100	385777	1323	problem	Distinction
100	23995020	701820	problem	Distinction
100	1048163	4880	course	Distinction
100	5111561	59423	problem	Distinction
100	572152	15960	problem	Distinction
100	529686	3723	sequential	Distinction
33.33	30002	322	chapter	Fail
100	6250480	84080	sequential	Fail
33.33	3564	81	sequential	Fail
33.33	1116	18	chapter	Fail
83.33	151650	1525	chapter	Fail
100	3816736	81543	course	Fail

Percent_grade	**SessionLength(sec)**	**NumEventsInSession**	**Module_type**	**FinalGrade**
66.67	143424	1044	chapter	Fail
100	21346	208	problem	Fail
83.33	241306	2886	sequential	Fail
66.67	440418	3465	chapter	Fail
33.33	18889	377	problem	Fail
50	542848	21152	chapter	Fail
100	43434	342	course	Fail
100	4344	912	chapter	Fail
66.67	284416	6226	video	Fail
100	2123990	13805	chapter	Fail
66.67	1824472	17992	chapter	Fail
100	958783	9457	chapter	Fail
100	295698	4131	sequential	Fail
16.67	14928	816	sequential	Fail
100	59416	3472	chapter	Fail
50	427175	11515	course	Fail
83.33	11229	456	chapter	Fail
50	2655	180	chapter	Fail
16.67	216961	2679	chapter	Fail
100	365296	12444	video	Fail
16.67	6180	12	course	Fail
100	41384	2128	sequential	Fail
100	144647	1265	sequential	Fail
16.67	1418784	14400	sequential	Fail
33.33	4432442	54733	sequential	Fail
100	2188571	72611	sequential	Fail
100	696444	7896	course	Fail
100	181560	3264	course	Fail
33.33	4944	12	problem	Fail
33.33	391923	14910	course	Fail
83.33	202554	2475	course	Fail
100	2527	95	problem	Fail
50	119102	1240	chapter	Fail
50	364584	20240	sequential	Fail

Percent_grade	**SessionLength(sec)**	**NumEventsInSession**	**Module_type**	**FinalGrade**
100	2785730	110334	problem	Fail
100	3109429	30876	chapter	Fail
100	1417409	30788	chapter	Fail
50	138402	1260	video	Fail
100	62792	3337	sequential	Fail
16.67	149616	5400	sequential	Fail
50	191777	3689	sequential	Fail
16.67	174100	1275	sequential	Fail
100	22661819	368406	course	Fail
80	1467461	16684	problem	Fail
100	6314625	303450	sequential	Fail
16.67	6461	221	sequential	Fail
16.67	9192	192	chapter	Fail
83.33	116028	704	sequential	Fail
25	3454	55	sequential	Fail
100	60	20	chapter	Fail
83.33	2493021	21150	sequential	Fail
16.67	1009935	44100	chapter	Fail
57.14	5120	200	course	Fail
100	1106370	32040	sequential	Fail
100	1962940	78470	chapter	Fail
100	56700	648	sequential	Fail
100	81980	10320	sequential	Fail
100	3502278	42042	course	Fail
16.67	10699	442	chapter	Fail
66.67	43280	2976	problem	Fail
75	4849234	358586	problem	Fail
66.67	452088	3744	chapter	Fail
40	9849	294	chapter	Fail
100	3821076	45528	sequential	Fail
33.33	806	117	chapter	Fail
100	3548748	38220	video	Fail
83.33	518630	6634	sequential	Fail
16.67	231390	1674	video	Fail

Percent_grade	SessionLength(sec)	NumEventsInSession	Module_type	FinalGrade
16.67	984	36	problem	Fail
100	141466	2340	video	Fail
80	319498	4692	problem	Fail
33.33	500066	5796	video	Fail
25	3216918	28290	sequential	Fail
16.67	720	12	chapter	Fail
100	7322861	80171	video	Fail
83.33	7570450	224480	video	Fail
66.67	25470	170	sequential	Fail
33.33	4090	80	course	Fail
100	578748	6528	sequential	Fail
50	32396	840	course	Fail
16.67	71324	902	problem	Fail
50	91050	1100	sequential	Fail
50	2450	40	sequential	Fail
100	81512	736	sequential	Fail
100	2130332	23316	sequential	Fail
83.33	201860	1040	sequential	Fail
100	5807680	140608	video	Fail
100	2384370	26136	video	Fail
100	1596627	20178	chapter	Fail
50	30	30	chapter	Fail
100	3660280	135148	sequential	Fail
83.33	1005990	16140	chapter	Fail
100	121550	1352	sequential	Fail
100	283775	1625	problem	Fail
100	626340	7290	chapter	Fail
66.67	572814	10956	chapter	Fail
83.33	39300	900	chapter	Fail
33.33	78360	552	problem	Fail
100	1928340	20340	chapter	Fail
100	173147	4313	chapter	Fail
100	285300	1830	sequential	Fail
100	4202373	54946	chapter	Fail

Percent_grade	**SessionLength(sec)**	**NumEventsInSession**	**Module_type**	**FinalGrade**
50	79995	1335	sequential	Fail
100	2557956	13668	problem	Fail
83.33	105	105	sequential	Fail
66.67	354970	12425	chapter	Fail
100	463250	53925	video	Fail
50	112800	975	chapter	Fail
83.33	3184510	54250	sequential	Fail
60	540040	23644	problem	Fail
100	1623812	54199	chapter	Fail
100	21830865	435540	video	Fail
100	1400059	13617	course	Fail
100	1012823	14104	chapter	Fail
33.33	71288	672	problem	Fail
66.67	550671	5973	chapter	Fail
50	480792	3000	problem	Fail
83.33	1964286	15225	chapter	Fail
100	411138	4329	chapter	Fail
33.33	553	21	sequential	Fail
83.33	264390	2310	sequential	Fail
83.33	56448	720	sequential	Fail
66.67	58604	504	chapter	Fail
83.33	1246350	11490	video	Fail
100	428329	8648	sequential	Fail
66.67	275010	1800	problem	Fail
16.67	121569	1386	chapter	Fail
40	169488	4136	sequential	Fail
33.33	8832	64	sequential	Fail
100	3027861	22359	problem	Fail
100	31977	759	problem	Fail
100	1069458	10608	chapter	Fail
100	601152	6324	problem	Fail
16.67	47376	392	sequential	Fail
28.57	4140	80	chapter	Fail
16.67	6	6	sequential	Fail

Percent_grade	SessionLength(sec)	NumEventsInSession	Module_type	FinalGrade
50	93090	551	course	Fail
16.67	29985	885	sequential	Fail
100	219852	1922	course	Fail
50	374500	3136	course	Fail
100	967681	22422	chapter	Fail
16.67	24427350	301800	video	Fail
100	504504	8400	course	Fail
16.67	99099	1419	sequential	Fail
100	70400	850	sequential	Fail
33.33	350433	3996	course	Fail
50	28025	228	chapter	Fail
100	1058080	17320	chapter	Fail
100	884442	5612	course	Fail
50	90416	1056	chapter	Fail
16.67	23208	204	sequential	Fail
66.67	129246	858	sequential	Fail
57.14	768876	11832	problem	Fail
100	449595	18810	chapter	Fail
100	579675	3717	problem	Fail
25	18584	207	problem	Fail
100	101711	775	sequential	Fail
75	151935	1435	chapter	Fail
100	475	114	problem	Fail
100	4009	152	course	Fail
100	26961	171	problem	Fail
100	1552680	12000	course	Merit
100	5428338	43962	chapter	Merit
100	1032953	11446	chapter	Merit
100	1342235	28560	problem	Merit
100	397660	3835	chapter	Merit
100	1390430	77010	problem	Merit
100	618688	3080	problem	Merit
100	283275	900	sequential	Merit
100	3355380	66780	problem	Pass

Percent _grade	SessionLength(sec)	NumEventsInSession	Module _type	FinalGrade
100	3495168	226329	problem	Pass
100	7892292	41412	chapter	Pass
100	545532	2496	problem	Pass
100	6141408	763680	video	Pass
100	2483299	65611	problem	Pass
100	302524	2067	problem	Pass
100	5708400	120800	problem	Pass
100	704047	12803	problem	Pass
100	53760	756	chapter	Pass
100	398060	1404	problem	Pass

Table B.1: Complete features used in Phase 1 to predict final grade

B.2 Adding Emotion/Sentiment Features to the Dataset in Phase 3

SessionLength(sec)	NumEventsInSession	percent_grade	down_count	up_count	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Finalgrade
43434	342	100	0	0	0	0	0	0	0	0	Fail
112800	975	50	0	0	0.1	0.1	0	0	0.1	0.1	Fail
119102	1240	50	0	0	0	0	0	0	0	0	Fail
119102	1240	50	0	1	0.1	0.1	0.1	1	0.3	0.1	Fail
958783	9457	100	0	0	0.1	0	0	0.1	0.3	0	Fail
1058080	17320	100	0	0	0.1	0.1	0	0.1	0	0	Fail
1623812	54199	100	0	0	0	0	0	0	0	0	Fail
1623812	54199	100	0	0	0	0	0	0	0	0	Fail
1623812	54199	100	0	0	0.1	0.1	0.1	0.1	0.1	0.1	Fail
1928340	20340	100	0	0	0.1	0	0.1	0.1	0.2	0	Fail
2123990	13805	100	0	1	0.1	0	0	0.1	0.1	0.1	Fail
2123990	13805	100	0	0	0	0	0	0	0	0	Fail
2123990	13805	100	0	0	0.1	0	0	1	0.1	0.1	Fail
2123990	13805	100	0	0	0.1	0.1	0.1	0	0.2	0	Fail
2123990	13805	100	0	0	0.1	0.1	0.2	0.3	0.3	0	Fail
2123990	13805	100	0	0	0	0	0	0.1	0	0	Fail
2384370	26136	100	0	0	0.1	0.4	0.1	0.4	0.2	0.4	Fail
2384370	26136	100	0	0	0	0	0	0.1	0.1	0	Fail
2629604	36076	100	0	0	0	0	0	1	0	0	Distinction
2629604	36076	100	0	0	0.2	0	0	0.2	0.1	0.2	Distinction

SessionLength(sec)	NumEventsInSession	percent_grade	down_count	up_count	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Finalgrade
2629604	36076	100	0	0	0	0	0	0.3	0.3	0	Distinction
3216918	28290	25	0	0	0	0	0	0.1	0.1	0	Fail
3216918	28290	25	0	0	0	0	0	0	0	0	Fail
3216918	28290	25	0	0	0.1	0	0.1	0.1	0.2	0	Fail
3216918	28290	25	0	0	0	0	0	0	0	0	Fail
3216918	28290	25	0	0	0.5	0.1	0.1	0.1	0.1	0	Fail
3216918	28290	25	0	0	0.1	0	0	0.1	0	0	Fail
3216918	28290	25	0	0	0.5	0	0	0	0.1	0	Fail
3816736	81543	100	0	0	0	0	0	0	0	0	Fail
3816736	81543	100	0	0	0.7	0.7	0.8	0.1	0.1	0	Fail
4873033	58916	100	0	0	0.7	0.7	0.8	0.2	0.3	0.1	Distinction
7570450	224480	83.33	0	3	0.1	0.1	0	0.1	0.1	0.1	Fail
7570450	224480	83.33	0	0	0.1	0.1	0.1	0.3	0.1	0.1	Fail
10445416	150388	100	0	3	0.1	0.1	0.1	0.1	0.1	0	Distinction
10445416	150388	100	0	0	0	0	0	0.1	0	0	Distinction
10445416	150388	100	0	2	0.1	0	0	0.3	0.3	0.1	Distinction
10445416	150388	100	0	0	0.1	0.1	0.1	0	0.1	0	Distinction
10445416	150388	100	0	0	0.1	0.3	0.1	1	0.2	1	Distinction
10445416	150388	100	0	0	0.1	0.2	0	0.1	0.2	0.1	Distinction
10445416	150388	100	0	1	0	0	0.1	0.5	0.3	0	Distinction

SessionLength(sec)	NumEventsInSession	percent_grade	down_count	up_count	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Finalgrade
11755891	200158	100	0	0	0.1	0.2	0	0.1	0.3	0.1	Distinction
11755891	200158	100	0	0	0	0	0.1	1	0.2	0	Distinction
11755891	200158	100	0	0	0	0	0	0.2	0	0	Distinction
11755891	200158	100	0	0	0	0	0	0.1	0	0	Distinction
11755891	200158	100	0	0	0.1	0	0	0.6	0.1	0.1	Distinction
11755891	200158	100	0	1	0.1	0	0.2	0.1	0.1	0.1	Distinction
11755891	200158	100	0	0	0.1	0	0	0.3	0.3	0	Distinction
11755891	200158	100	0	2	0	0	0	0.1	0.3	0	Distinction
11755891	200158	100	0	0	0	0	0	1	0	0	Distinction
11755891	200158	100	0	0	0.1	0.1	0.1	0.1	0.2	0.1	Distinction
11755891	200158	100	0	0	0	0	0	0.1	0.2	0	Distinction
11755891	200158	100	0	0	0.1	0	0	0.1	0.1	0.1	Distinction
11755891	200158	100	0	1	0.1	0.8	0.2	0.1	0.1	0.1	Distinction
11755891	200158	100	0	0	0.1	0.1	0.4	0.4	0.4	0	Distinction
11755891	200158	100	0	0	0	0	0	0.1	0.1	0	Distinction
11755891	200158	100	0	0	0.1	0	0	0.1	0.2	0	Distinction
11755891	200158	100	0	0	0.1	0	0	0	0	0	Distinction
11755891	200158	100	0	1	0.1	0.1	0.1	0.1	0.2	0.1	Distinction
11755891	200158	100	0	0	0	0	0	0.1	0.4	0	Distinction
11755891	200158	100	0	1	0	0	0	0	0	0	Distinction

SessionLength(sec)	NumEventsInSession	percent_grade	down_count	up_count	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Finalgrade
13283456	235264	100	0	0	0	0	0	0.3	0.3	0	Distinction
13283456	235264	100	0	0	0	0	0	1	0	0	Distinction
19252548	410400	100	0	0	0.1	0.1	0	0.1	0.1	0.1	Distinction
19252548	410400	100	0	2	0	0	0	0.1	0.8	0	Distinction
19252548	410400	100	0	0	0.2	0.2	0.4	0.4	0.4	0	Distinction
25570846	823669	100	0	1	0	0	0.1	0.1	0.2	0	Distinction
25570846	823669	100	0	0	0.1	0	0	0.2	0.2	0.1	Distinction
25570846	823669	100	0	0	0.1	0.1	0.2	0	0.2	0	Distinction
25570846	823669	100	0	0	0.1	0.1	0.1	0.3	0.3	0.1	Distinction
25570846	823669	100	0	0	0.1	0	0	0.2	0.1	0	Distinction
25570846	823669	100	0	0	0	0	0.1	0.1	0	0	Distinction

Table B.2: Adding emotion/sentiment features to the dataset in Phase 3

B.3 Summary of the One-way ANOVA Test for the Six Emotions

t-Test: Two-Sample Assuming Unequal Variances

	Happiness	**Disgust**
Mean	0.216174563	0.067559099
Variance	0.08342298	0.025039179
Observations	71	71
Hypothesized Mean Difference	0	
df	109	
t Stat	3.802367197	
P(T<=t) one-tail	0.000118345	
t Critical one-tail	1.658953458	
P(T<=t) two-tail	0.000236689	
t Critical two-tail	1.98196749	

	Happiness	**Fear**
Mean	0.216174563	0.072539592
Variance	0.08342298	0.022878904
Observations	71	71
Hypothesized Mean Difference	0	
df	106	
t Stat	3.712093327	
P(T<=t) one-tail	0.000164783	
t Critical one-tail	1.659356034	
P(T<=t) two-tail	0.000329565	
t Critical two-tail	1.982597262	

	Happiness	**Anger**
Mean	0.216174563	0.079492676
Variance	0.08342298	0.020082607
Observations	71	71
Hypothesized Mean Difference	0	
df	102	
t Stat	3.579795711	
P(T<=t) one-tail	0.000264135	
t Critical one-tail	1.659929976	
P(T<=t) two-tail	0.000528271	
t Critical two-tail	1.983495259	

	Happiness	**Sadness**
Mean	0.216174563	0.146267141
Variance	0.08342298	0.020628699
Observations	71	71
Hypothesized Mean Difference	0	
df	103	
t Stat	1.826114131	
P(T<=t) one-tail	0.03536497	
t Critical one-tail	1.659782273	
P(T<=t) two-tail	0.070729939	
t Critical two-tail	1.983264145	

	Happiness	**Surprise**
Mean	0.216174563	0.043796139
Variance	0.08342298	0.01607516
Observations	71	71
Hypothesized Mean Difference	0	
df	96	
t Stat	4.609909599	
P(T<=t) one-tail	6.20498E-06	
t Critical one-tail	1.66088144	
P(T<=t) two-tail	1.241E-05	
t Critical two-tail	1.984984312	

	Disgust	**Fear**
Mean	0.067559099	0.072539592
Variance	0.025039179	0.022878904
Observations	71	71
Hypothesized Mean Difference	0	
df	140	
t Stat	-0.191713098	
P(T<=t) one-tail	0.424122361	
t Critical one-tail	1.655810511	
P(T<=t) two-tail	0.848244723	
t Critical two-tail	1.97705372	

	Disgust	**Anger**
Mean	0.067559099	0.079492676
Variance	0.025039179	0.020082607
Observations	71	71
Hypothesized Mean Difference	0	
df	138	
t Stat	-0.473376491	
P(T<=t) one-tail	0.318346282	
t Critical one-tail	1.655970382	
P(T<=t) two-tail	0.636692564	
t Critical two-tail	1.977303542	

	Disgust	**Sadness**
Mean	0.067559099	0.146267141
Variance	0.025039179	0.020628699
Observations	71	71
Hypothesized Mean Difference	0	
df	139	
t Stat	-3.103436516	
P(T<=t) one-tail	0.001158951	
t Critical one-tail	1.655889868	
P(T<=t) two-tail	0.002317902	
t Critical two-tail	1.977177724	

	Disgust	**Surprise**
Mean	0.067559099	0.043796139
Variance	0.025039179	0.01607516
Observations	71	71
Hypothesized Mean Difference	0	
df	134	
t Stat	0.990182968	
P(T<=t) one-tail	0.161934658	
t Critical one-tail	1.656304542	
P(T<=t) two-tail	0.323869317	
t Critical two-tail	1.977825758	

	Fear	**Anger**
Mean	0.072539592	0.079492676
Variance	0.022878904	0.020082607
Observations	71	71
Hypothesized Mean Difference	0	
df	139	
t Stat	-0.282661661	
P(T<=t) one-tail	0.388928345	
t Critical one-tail	1.655889868	
P(T<=t) two-tail	0.77785669	
t Critical two-tail	1.977177724	

	Fear	**Sadness**
Mean	0.072539592	0.146267141
Variance	0.022878904	0.020628699
Observations	71	71
Hypothesized Mean Difference	0	
df	140	
t Stat	-2.97835453	
P(T<=t) one-tail	0.001708363	
t Critical one-tail	1.655810511	
P(T<=t) two-tail	0.003416725	
t Critical two-tail	1.97705372	

	Fear	**Surprise**
Mean	0.072539592	0.043796139
Variance	0.022878904	0.01607516
Observations	71	71
Hypothesized Mean Difference	0	
df	136	
t Stat	1.230665107	
P(T<=t) one-tail	0.110286189	
t Critical one-tail	1.656134988	
P(T<=t) two-tail	0.220572378	
t Critical two-tail	1.977560777	

	Anger	**Sadness**
Mean	0.079492676	0.146267141
Variance	0.020082607	0.020628699
Observations	71	71
Hypothesized Mean Difference	0	
df	140	
t Stat	-2.788573329	
P(T<=t) one-tail	0.003015524	
t Critical one-tail	1.655810511	
P(T<=t) two-tail	0.006031047	
t Critical two-tail	1.97705372	

	Anger	**Surprise**
Mean	0.079492676	0.043796139
Variance	0.020082607	0.01607516
Observations	71	71
Hypothesized Mean Difference	0	
df	139	
t Stat	1.586716916	
P(T<=t) one-tail	0.057424523	
t Critical one-tail	1.655889868	
P(T<=t) two-tail	0.114849047	
t Critical two-tail	1.977177724	

	Sadness	**Surprise**
Mean	0.146267141	0.043796139
Variance	0.020628699	0.01607516
Observations	71	71
Hypothesized Mean Difference	0	
df	138	
t Stat	4.520632599	
P(T<=t) one-tail	6.57245E-06	
t Critical one-tail	1.655970382	
P(T<=t) two-tail	1.31449E-05	
t Critical two-tail	1.977303542	

Appendix C

Complete Results for Experiment 3

C.1 Experiment 3:Complete Results of Emotion Extraction and Writing Style Features for client's Sessions

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Ability	Client004_a	sad	20	0.334792123	0.411255411
Relationships	Client004_b	sad	40	0.3	0.495934959
Relationships	Client004_c	sad	33	0.310679612	0.506410256
Relationships	Client004_d	sad	22	0.426035503	0.41509434
Relationships	Client004_e	sad	9	0.381322957	0.397590361
Behaviour	Client004_f	sad	15	0.388779528	0.352601156
Behaviour	Client004_g	peaceful	21	0.359242325	0.473170732
Behaviour	Client004_h	indifferent	11	0.355733662	0.497409326
Ability	Client004_i	sad	177	0.200187091	0.450715421
Ability	Client011_a	Friendly	945	0.112697198	0.824940048
Ability	Client011_b	Friendly	146	0.229634672	0.83203125
Relationships	Client011_c	Friendly	64	0.29015919	0.833333333
Relationships	Client011_d	Friendly	126	0.260787992	0.884615385
Relationships	Client011_e	Friendly	143	0.256395178	0.752941176
Ability	Client011_f	peaceful	67	0.238571815	0.90797546

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Behaviour	Client011_g	peaceful	54	0.199496855	0.79342723
Culture	Client011_h	Friendly	113	0.22172619	0.790598291
Relationships	Client011_i	Friendly	142	0.265526553	0.820512821
Relationships	Client016_a	peaceful	41	0.315902579	0.669421488
Relationships	Client016_b	Friendly	14	0.433849821	0.627118644
Relationships	Client016_c	Friendly	46	0.343678686	0.645714286
Relationships	Client016_d	peaceful	49	0.382887189	0.725352113
Relationships	Client016_e	peaceful	41	0.415368082	0.687022901
Development	Client016_f	peaceful	33	0.364278507	0.78030303
Relationships	Client016_g	peaceful	36	0.278882576	0.74796748
Behaviour	Client016_h	peaceful	29	0.322977346	0.666666667
Behaviour	Client016_i	peaceful	32	0.304277206	0.833333333
Development	Client018_a	sad	73	0.248955224	0.699152542
Relationships	Client018_b	peaceful	51	0.302836596	0.716666667
Behaviour	Client018_c	sad	57	0.320996979	0.75
Ability	Client018_d	sad	39	0.292576419	0.668639053
Personality	Client018_e	peaceful	63	0.274716029	0.624277457
Relationships	Client018_f	sad	78	0.301617149	0.71
Personality	Client018_g	sad	42	0.361516035	0.823529412
Relationships	Client018_h	peaceful	59	0.241905471	0.628318584
Ability	Client018_i	peaceful	71	0.344947735	0.841269841
Behaviour	Client018_j	peaceful	36	0.303336704	0.771929825
Personality	Client018_k	peaceful	42	0.293458619	0.762295082
Relationships	Client019_a	sad	42	0.335827099	0.480314961
Culture	Client019_b	peaceful	19	0.427258806	0.341463415
Relationships	Client019_c	peaceful	4	0.409244645	0.290909091
Relationships	Client019_d	sad	6	0.295031056	0.297029703
Relationships	Client019_e	sad	15	0.379310345	0.362068966
Relationships	Client019_f	peaceful	4	0.34430727	0.326388889
Development	Client019_g	sad	21	0.275030902	0.38410596
Behaviour	Client019_h	peaceful	14	0.368909513	0.339449541
Ability	Client019_i	peaceful	8	0.371505861	0.364341085
Development	Client019_j	peaceful	11	0.327545383	0.415300546
Culture	Client019_k	peaceful	18	0.380355277	0.515789474

Topic	Session_name	Emotion	Emo_value	Richness	POS
Relationships	Client031_a	Friendly	26	0.312348668	0.614173228
Ability	Client031_b	Friendly	47	0.281586022	0.552238806
Health	Client031_c	Friendly	38	0.331909701	0.601851852
Relationships	Client031_d	admire	43	0.331532748	0.53125
Ability	Client031_e	Friendly	46	0.379844961	0.601851852
Relationships	Client031_f	peaceful	20	0.371120108	0.597402597
Health	Client031_g	Friendly	31	0.329383886	0.587719298
Culture	Client031_h	peaceful	17	0.299207398	0.577586207
Behaviour	Client031_i	peaceful	35	0.324137931	0.654411765
Personality	Client031_j	sad	15	0.413793103	0.432432432
Relationships	Client031_k	Friendly	27	0.334502104	0.463414634
Ability	Client031_l	peaceful	20	0.322201608	0.406779661
Ability	Client031_m	Friendly	17	0.29956427	0.436781609
Relationships	Client031_n	peaceful	16	0.337099812	0.5
Relationships	Client032_a	sad	9	0.284776903	0.280373832
Relationships	Client032_b	sad	18	0.244795406	0.323383085
Behaviour	Client032_c	peaceful	18	0.303212851	0.343065693
Relationships	Client032_d	alarmed	15	0.293440736	0.3984375
Relationships	Client032_e	peaceful	22	0.259813084	0.369747899
Behaviour	Client032_f	admire	10	0.322368421	0.302083333
Relationships	Client032_g	alarmed	11	0.284782609	0.388349515
Relationships	Client032_h	peaceful	9	0.288924559	0.443708609
Ability	Client032_i	peaceful	29	0.263542301	0.481481481
Relationships	Client032_j	furious	35	0.3175	0.567567568
Relationships	Client032_k	sad	22	0.297413793	0.555555556
Ability	Client032_l	sad	44	0.287434161	0.451428571
Personality	Client032_m	sad	26	0.394409938	0.488372093
Relationships	Client034_a	Friendly	32	0.391280353	0.636363636
Relationships	Client034_b	peaceful	25	0.354561102	0.756097561
Relationships	Client034_c	peaceful	25	0.361477573	0.75308642
Relationships	Client034_d	peaceful	35	0.366103203	0.624413146
Relationships	Client034_e	alert	29	0.31504639	0.503401361
Relationships	Client034_f	peaceful	40	0.337980566	0.390374332
Relationships	Client034_g	alert	32	0.308679707	0.471428571

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Relationships	Client034_h	admire	37	0.333790267	0.488888889
Relationships	Client034_i	peaceful	1	0.739130435	0.4
Relationships	Client034_j	peaceful	51	0.343634116	0.632653061
Development	Client034_k	alert	34	0.393956044	0.823529412
Personality	Client034_l	peaceful	33	0.359060403	0.817204301
Relationships	Client038_a	alarmed	23	0.39466896	0.528089888
Development	Client038_b	sad	19	0.388888889	0.523809524
Behaviour	Client038_c	peaceful	20	0.374670185	0.602564103
Behaviour	Client105_a	peaceful	27	0.357180157	0.569444444
Development	Client105_b	peaceful	26	0.350310559	0.548611111
Behaviour	Client105_c	sad	29	0.443466486	0.657142857
Behaviour	Client105_d	peaceful	31	0.406806283	0.714285714
Health	Client105_e	peaceful	20	0.403183024	0.6625
Relationships	Client105_f	peaceful	36	0.36983842	0.597122302
Behaviour	Client105_g	peaceful	23	0.393772894	0.768518519
Behaviour	Client105_h	peaceful	34	0.364741641	0.656
Health	Client105_i	sad	16	0.387968614	0.495867769
Relationships	Client105_j	peaceful	17	0.386391252	0.616541353
Relationships	Client105_k	Friendly	18	0.387096774	0.593984962
Relationships	Client110_a	Friendly	19	0.393801965	0.589473684
Development	Client110_b	peaceful	50	0.317160827	0.706422018
Relationships	Client110_c	peaceful	39	0.404371585	0.717647059
Behaviour	Client110_d	peaceful	20	0.361350575	0.653333333
Relationships	Client110_e	alert	37	0.300068353	0.465648855
Relationships	Client110_f	peaceful	34	0.336470588	0.504587156
Relationships	Client110_g	peaceful	31	0.332368548	0.559748428
Relationships	Client110_h	peaceful	26	0.37446198	0.488888889
Relationships	Client110_i	peaceful	43	0.330994898	0.547368421
Relationships	Client110_j	peaceful	48	0.298520453	0.551971326
Culture	Client110_k	peaceful	16	0.412063953	0.533742331
Relationships	Client110_l	peaceful	45	0.30418251	0.6125
Relationships	Client110_m	peaceful	30	0.347639485	0.617283951
Ability	Client110_n	furious	24	0.349573079	0.61682243
Development	Client110_o	cheerful	27	0.340224454	0.606060606

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Relationships	Client110_p	peaceful	19	0.33354232	0.555555556
Development	Client110_q	peaceful	34	0.283119658	0.569230769
Relationships	Client110_r	admire	42	0.283105023	0.489051095
Ability	Client112_a	peaceful	49	0.300936246	0.600877193
Ability	Client112_b	sad	14	0.397034596	0.595744681
Relationships	Client112_c	peaceful	22	0.263236763	0.514851485
Development	Client112_d	Friendly	19	0.340213049	0.622377622
Behaviour	Client112_e	peaceful	19	0.312451057	0.509259259
Relationships	Client112_f	alarmed	13	0.312989045	0.516393443
Relationships	Client112_g	sad	23	0.358807083	0.422680412
Ability	Client112_h	cheerful	55	0.346740638	0.520833333
Relationships	Client112_i	sad	18	0.297242771	0.514851485
Relationships	Client112_j	sad	42	0.36770428	0.480769231
Relationships	Client123_a	admire	14	0.405673759	0.509803922
Development	Client123_b	sad	26	0.31024735	0.330218069
Ability	Client123_c	sad	9	0.570512821	0.366666667
Relationships	Client123_d	peaceful	15	0.365079365	0.598684211
Culture	Client123_e	peaceful	20	0.321263482	0.4375
Personality	Client123_f	peaceful	25	0.312462731	0.512345679
Relationships	Client123_g	sad	10	0.327257664	0.595041322
Relationships	Client123_h	indifferent	16	0.336612022	0.495049505
Relationships	Client123_i	cheerful	23	0.36154289	0.569767442
Behaviour	Client123_j	peaceful	16	0.371518987	0.563492063
Health	Client123_k	sad	25	0.406490649	0.540740741
Development	Client123_l	sad	36	0.397268777	0.515957447
Development	Client123_m	sad	40	0.356772334	0.521126761
Relationships	Client123_n	Friendly	19	0.338765009	0.596330275
Relationships	Client123_o	sad	25	0.35765838	0.593495935
Behaviour	Client123_p	peaceful	23	0.322461538	0.466216216
Relationships	Client124_a	peaceful	20	0.425855513	0.680555556
Behaviour	Client124_b	Friendly	23	0.327403643	0.602272727
Relationships	Client124_c	alert	40	0.295472287	0.620967742
Relationships	Client124_d	sad	38	0.338655056	0.728
Relationships	Client124_e	alert	11	0.356092437	0.613636364

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Behaviour	Client124_f	peaceful	20	0.280509219	0.60952381
Behaviour	Client124_g	peaceful	14	0.290273556	0.462809917
Personality	Client124_h	peaceful	24	0.33464761	0.720338983
Behaviour	Client124_i	alert	39	0.345953003	0.653846154
Personality	Client124_j	alert	30	0.336030462	0.513368984
Personality	Client124_k	peaceful	32	0.363031234	0.762295082
Relationships	Client124_l	peaceful	29	0.329277865	0.694610778
Relationships	Client124_m	peaceful	36	0.308608059	0.582677165
Behaviour	Client124_n	peaceful	34	0.352592593	0.643564356
Relationships	Client124_o	alert	29	0.362733645	0.761904762
Ability	Client124_p	alert	35	0.3125	0.642384106
Development	Client124_q	peaceful	34	0.344618056	0.760416667
Personality	Client124_r	peaceful	34	0.300988649	0.790697674
Relationships	Client124_s	sad	30	0.330982368	0.767123288
Behaviour	Client124_t	sad	7	0.449324324	0.5
Personality	Client124_u	peaceful	36	0.34494382	0.694214876
Behaviour	Client124_v	peaceful	18	0.390597795	0.621848739
Ability	Client137_a	Friendly	42	0.41560219	0.603305785
Relationships	Client137_b	alert	30	0.366717248	0.538461538
Relationships	Client137_c	Friendly	40	0.356431701	0.442622951
Ability	Client137_d	admire	27	0.383789587	0.492753623
Development	Client137_e	peaceful	26	0.394554191	0.559322034
Personality	Client137_f	alert	31	0.268266085	0.378723404
Relationships	Client137_g	alert	22	0.474116162	0.550561798
Culture	Client137_h	admire	19	0.377310062	0.657894737
Relationships	Client137_i	sad	25	0.436139332	0.580645161
Development	Client137_j	peaceful	30	0.397735315	0.521276596
Behaviour	Client137_k	peaceful	26	0.425952045	0.632911392
Development	Client137_l	peaceful	14	0.46856465	0.507936508
Behaviour	Client137_m	sad	33	0.376584188	0.358585859
Personality	Client137_n	sad	18	0.429725363	0.401234568
Relationships	Client137_o	sad	20	0.429721816	0.5
Behaviour	Client201_a	furious	25	0.257537688	0.613861386
Ability	Client201_b	peaceful	12	0.303343949	0.64516129

Topic	Session_name	Emotion	Emo_value	Richness	POS
Personality	Client201_c	peaceful	11	0.31779661	0.569620253
Development	Client202_a	peaceful	38	0.357730619	0.573770492
Relationships	Client202_b	sad	36	0.386865059	0.670454545
Behaviour	Client202_c	peaceful	39	0.366838891	0.769230769
Health	Client202_d	peaceful	28	0.39912759	0.72826087
Behaviour	Client202_e	peaceful	22	0.35579782	0.698795181
Behaviour	Client203_a	Friendly	24	0.362644416	0.733333333
Development	Client203_b	peaceful	20	0.366477273	0.442307692
Relationships	Client204_a	Friendly	41	0.309034268	0.579439252
Behaviour	Client204_b	Friendly	30	0.377850163	0.392307692
Relationships	Client204_c	sad	40	0.280294451	0.341968912
Behaviour	Client205_a	peaceful	31	0.307232516	0.47826087
Relationships	Client205_b	peaceful	43	0.309197652	0.510638298
Behaviour	Client205_c	admire	14	0.329054054	0.578431373
Relationships	Client205_d	peaceful	23	0.275242047	0.41221374
Ability	Client205_e	peaceful	26	0.297702298	0.461538462
Relationships	Client205_f	peaceful	10	0.388888889	0.5
Development	Client205_g	peaceful	23	0.345844504	0.411764706
Behaviour	Client205_h	furious	29	0.283070596	0.45
Behaviour	Client206_a	sad	27	0.308379413	0.462222222
Behaviour	Client206_b	peaceful	31	0.336706015	0.495238095
Development	Client207_a	peaceful	13	0.304971319	0.415254237
Development	Client207_b	sad	18	0.307592472	0.653061224
Development	Client207_c	sad	37	0.267142291	0.471875
Relationships	Client208_a	peaceful	24	0.269249632	0.38317757
Health	Client208_b	furious	39	0.276225619	0.431034483
Relationships	Client208_c	sad	33	0.356108859	0.6
Relationships	Client209_a	sad	17	0.404009252	0.553398058
Relationships	Client209_b	sad	21	0.427707199	0.448979592
Relationships	Client209_c	peaceful	48	0.386445566	0.491803279
Behaviour	Client210_a	sad	16	0.453102453	0.568965517
Behaviour	Client210_b	Friendly	13	0.423390752	0.64556962
Culture	Client210_c	peaceful	13	0.337059329	0.569620253
Relationships	Client211_a	Friendly	26	0.340884574	0.648648649

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Relationships	Client211_b	sad	17	0.399725275	0.534883721
Behaviour	Client211_c	sad	21	0.340381992	0.587628866
Relationships	Client211_d	peaceful	23	0.330255682	0.515151515
Relationships	Client212_a	peaceful	17	0.465777778	0.466165414
Behaviour	Client212_b	peaceful	32	0.396179914	0.403846154
Ability	Client212_c	sad	29	0.397427653	0.366666667
Relationships	Client212_d	peaceful	29	0.400294334	0.389830508
Relationships	Client213_a	indifferent	17	0.252705628	0.519553073
Relationships	Client214_a	Friendly	34	0.310596833	0.361867704
Development	Client214_b	peaceful	28	0.322508399	0.476744186
Health	Client214_c	peaceful	25	0.341671751	0.622516556
Relationships	Client215_a	peaceful	31	0.375154512	0.745762712
Behaviour	Client215_b	peaceful	37	0.325409403	0.407035176
Relationships	Client216_a	peaceful	41	0.314055637	0.58974359
Development	Client216_b	sad	36	0.328918322	0.655172414
Relationships	Client216_c	sad	53	0.335791265	0.669064748
Relationships	Client216_d	sad	19	0.404586405	0.705882353
Relationships	Client217_a	peaceful	22	0.4646098	0.648648649
Development	Client217_b	peaceful	22	0.475867909	0.476190476
Development	Client217_c	admire	14	0.435120435	0.532467532
Relationships	Client217_d	sad	20	0.384375	0.457831325
Development	Client217_e	peaceful	18	0.500806452	0.622222222
Relationships	Client218_a	peaceful	6	0.29455081	0.438095238
Relationships	Client219_a	peaceful	38	0.272438443	0.764285714
Behaviour	Client219_b	sad	27	0.341409692	0.746987952
Behaviour	Client219_c	peaceful	43	0.292220114	0.652173913
Relationships	Client219_d	peaceful	15	0.416216216	0.594594595
Relationships	Client219_e	sad	23	0.331550802	0.488888889
Relationships	Client219_f	peaceful	50	0.277823241	0.598039216
Relationships	Client219_g	sad	46	0.368866328	0.846153846
Relationships	Client220_a	Friendly	15	0.314572864	0.33974359
Behaviour	Client220_b	brave	12	0.3372835	0.345323741
Development	Client220_c	Friendly	27	0.352900552	0.365217391
Development	Client221_a	sad	33	0.346291332	0.455696203

Topic	Session_name	Emotion	Emo_value	Richness	POS
Relationships	Client221_b	peaceful	32	0.359960552	0.478873239
Ability	Client221_c	peaceful	25	0.32792887	0.453900709
Behaviour	Client221_d	admire	19	0.384848485	0.489583333
Relationships	Client222_a	alert	8	0.353794643	0.530864198
Relationships	Client222_b	peaceful	15	0.414211438	0.605263158
Behaviour	Client222_c	peaceful	13	0.393491124	0.56626506
Relationships	Client222_d	peaceful	8	0.313686968	0.45045045
Relationships	Client222_e	admire	12	0.326963907	0.456790123
Ability	Client223_a	sad	21	0.377037037	0.541176471
Relationships	Client223_b	peaceful	22	0.363752393	0.438461538
Relationships	Client224_a	sad	19	0.436619718	0.401315789
Relationships	Client224_b	peaceful	13	0.431686047	0.446428571
Relationships	Client224_c	sad	34	0.370707779	0.455357143
Relationships	Client224_d	peaceful	30	0.343480467	0.472049689
Relationships	Client224_e	peaceful	39	0.299924357	0.494505495
Ability	Client224_f	peaceful	19	0.421180275	0.639344262
Relationships	Client224_g	peaceful	25	0.421917808	0.481012658
Development	Client224_h	peaceful	18	0.429525223	0.589041096
Ability	Client224_i	peaceful	26	0.400259067	0.409090909
Relationships	Client224_j	peaceful	18	0.387011266	0.419354839
Relationships	Client224_k	peaceful	17	0.409556314	0.405797101
Relationships	Client224_l	peaceful	10	0.47640118	0.636363636
Ability	Client225_a	peaceful	24	0.366093366	0.636363636
Relationships	Client225_b	Friendly	10	0.383079848	0.666666667
Ability	Client225_c	peaceful	10	0.382497542	0.56626506
Relationships	Client225_d	peaceful	21	0.346676737	0.569444444
Relationships	Client226_a	peaceful	35	0.432713755	0.507317073
Relationships	Client226_b	peaceful	28	0.322473771	0.350148368
Relationships	Client226_c	sad	20	0.328919861	0.45
Relationships	Client227_a	sad	44	0.314881669	0.610526316
Behaviour	Client227_b	sad	14	0.323395982	0.557377049
Behaviour	Client228_a	sad	16	0.356620634	0.548780488
Relationships	Client228_b	peaceful	30	0.412144703	0.580246914
Ability	Client228_c	alert	11	0.287679426	0.444444444

Topic	Session_name	Emotion	Emo_value	Richness	POS
Relationships	Client229_a	cheerful	33	0.377348066	0.65
Relationships	Client229_b	peaceful	19	0.332947307	0.577235772
Relationships	Client230_a	peaceful	21	0.345794393	0.582089552
Development	Client230_b	peaceful	12	0.317152104	0.584415584
Ability	Client303_a	peaceful	5	0.614583333	0.29787234
Behaviour	Client303_b	peaceful	17	0.423452769	0.391304348
Ability	Client304_a	peaceful	13	0.484375	0.537037037
Behaviour	Client305_a	peaceful	8	0.505102041	0.310344828
Development	Client305_b	alert	5	0.554572271	0.229885057
Ability	Client307_a	peaceful	25	0.344883159	0.577981651
Behaviour	Client309_a	sad	5	0.586330935	0.309859155
Behaviour	Client310_a	peaceful	11	0.443946188	0.380530973
Behaviour	Client401_a	sad	38	0.38215103	0.516666667
Relationships	Client401_b	sad	21	0.410046729	0.389473684
Ability	Client402_a	sad	45	0.328815557	0.412408759
Relationships	Client402_b	peaceful	41	0.295964126	0.453781513
Personality	Client402_c	peaceful	35	0.242072963	0.501805054
Relationships	Client402_d	sad	24	0.305209513	0.304635762
Development	Client402_e	sad	56	0.269808917	0.537931034
Development	Client403_a	sad	20	0.450155763	0.526315789
Relationships	Client403_b	peaceful	16	0.468119451	0.557142857
Relationships	Client403_c	peaceful	28	0.497933884	0.671052632
Relationships	Client403_d	peaceful	17	0.539976825	0.518987342
Development	Client404_a	sad	25	0.406538139	0.708333333
Relationships	Client405_a	sad	45	0.31047266	0.482014388
Culture	Client405_b	peaceful	31	0.375753516	0.544554455
Culture	Client405_c	Friendly	78	0.302812071	0.643835616
Relationships	Client405_d	sad	37	0.3315	0.598214286
Relationships	Client405_e	peaceful	37	0.310322156	0.509803922
Relationships	Client405_f	Friendly	31	0.363319386	0.394736842
Ability	Client406_a	brave	6	0.545112782	0.666666667
Behaviour	Client406_b	sad	13	0.544247788	0.576923077
Development	Client406_c	alert	20	0.467402207	0.458823529
Behaviour	Client406_d	peaceful	17	0.442326981	0.346938776

Topic	Session_name	Emotion	Emo_value	Richness	POS
Relationships	Client406_e	admire	17	0.438672439	0.285714286
Behaviour	Client406_f	peaceful	17	0.46474359	0.465116279
Development	Client406_g	peaceful	20	0.4	0.355932203
Ability	Client406_h	sad	16	0.534351145	0.617647059
Development	Client406_i	peaceful	16	0.43598234	0.580246914
Development	Client406_j	peaceful	12	0.445595855	0.402439024
Behaviour	Client406_k	sad	15	0.502559727	0.436363636
Development	Client406_l	sad	10	0.506010929	0.6875
Development	Client406_m	sad	10	0.515083799	0.735294118
Behaviour	Client406_n	admire	10	0.526636225	0.406779661
Ability	Client406_o	admire	6	0.478079332	0.483870968
Development	Client406_p	peaceful	14	0.465250965	0.710526316
Development	Client407_a	peaceful	19	0.448875256	0.62295082
Relationships	Client408_a	peaceful	43	0.381120944	0.563106796
Development	Client408_b	peaceful	31	0.370480142	0.672131148
Ability	Client409_a	peaceful	38	0.479733481	0.622222222
Relationships	Client409_b	alert	12	0.496606335	0.642857143
Relationships	Client409_c	peaceful	9	0.394823789	0.652777778
Development	Client409_d	peaceful	19	0.427312775	0.702702703
Development	Client409_e	peaceful	22	0.473648649	0.763157895
Development	Client409_f	peaceful	16	0.403138794	0.627906977
Behaviour	Client410_a	cheerful	33	0.452418097	0.666666667
Development	Client410_b	sad	44	0.503196347	0.6
Personality	Client411_a	sad	33	0.365243004	0.666666667
Ability	Client411_b	peaceful	24	0.40148448	0.565656566
Behaviour	Client412_a	sad	11	0.432286024	0.25877193
Ability	Client413_b	Not available	0	0.611111111	0
Ability	Client413_c	sad	26	0.398799314	0.708333333
Development	Client414_a	Friendly	6	0.656441718	0.470588235
Behaviour	Client414_b	sad	10	0.604	0.32
Behaviour	Client414_c		0	0.555555556	0.4
Ability	Client414_d	peaceful	10	0.48177496	0.557377049
Ability	Client416_a	peaceful	10	0.48177496	0.557377049
Ability	Client416_b	sad	10	0.604	0.32

Topic	**Session_name**	**Emotion**	**Emo_value**	**Richness**	**POS**
Development	Client416_i	Friendly	6	0.656441718	0.470588235
Behaviour	Client416_j	sad	10	0.604	0.32
Behaviour	Client416_k	Not available	0	0.555555556	0.4
Behaviour	Client417_a	sad	8	0.325319309	0.412698413
Ability	Client417_b	peaceful	20	0.332357247	0.402985075
Personality	Client417_c	Friendly	15	0.373919874	0.414285714
Relationships	Client417_d	Friendly	32	0.309360731	0.422222222
Personality	Client417_e	peaceful	13	0.327993255	0.482352941
Relationships	Client417_f	Friendly	13	0.327402135	0.377192982
Relationships	Client417_g	peaceful	12	0.300097752	0.363636364
Ability	Client417_h	sad	25	0.322033898	0.721518987
Ability	Client417_i	peaceful	14	0.354585153	0.730769231
Development	Client417_j	peaceful	28	0.31714876	0.644067797
Development	Client417_k	peaceful	20	0.263565891	0.453271028
Ability	Client417_l	peaceful	22	0.284074605	0.383763838
Relationships	Client417_m	Friendly	19	0.251097454	0.42739726
Development	Client417_n	peaceful	37	0.252833908	0.399305556
Personality	Client418_a	sad	63	0.406311637	0.597560976
Relationships	Client418_b	sad	76	0.345837616	0.567010309
Ability	Client418_c	Friendly	14	0.4137577	0.508928571
Ability	Client418_d	sad	27	0.369318182	0.603773585
Relationships	Client418_e	sad	39	0.344914719	0.613138686
Culture	Client419_a	admire	10	0.469333333	0.535714286
Relationships	Client419_b	peaceful	17	0.420054201	0.591549296
Ability	Client419_c	cheerful	9	0.451476793	0.618181818
Personality	Client419_d	alert	19	0.373376623	0.392857143
Relationships	Client419_e	peaceful	9	0.439335888	0.547945205
Behaviour	Client419_f	admire	14	0.38121118	0.31547619
Relationships	Client419_g	peaceful	14	0.459057072	0.474358974
Relationships	Client419_h	alert	12	0.451318458	0.561403509
Relationships	Client419_i	peaceful	12	0.493243243	0.5
Behaviour	Client419_j	peaceful	11	0.440934066	0.568965517
Relationships	Client419_k	peaceful	15	0.430366492	0.634615385
Ability	Client419_l	peaceful	14	0.408921933	0.481132075

Topic	Session_name	Emotion	Emo_value	Richness	POS
Relationships	Client419_m	sad	15	0.47873633	0.404494382
Ability	Client419_n	alert	8	0.487323944	0.452830189
Relationships	Client419_o	admire	11	0.473895582	0.596491228
Relationships	Client419_p	alert	12	0.426253687	0.5
Relationships	Client419_q	peaceful	10	0.432	0.492753623
Ability	Client419_r	peaceful	16	0.358536585	0.434210526
Relationships	Client420_a	peaceful	20	0.311320755	0.642105263
Relationships	Client420_b	sad	64	0.316637881	0.736842105
Relationships	Client420_c	Friendly	10	0.336775218	0.608695652
Behaviour	Client420_d	peaceful	27	0.299863388	0.557894737
Behaviour	Client420_e	sad	38	0.335623679	0.579617834
Behaviour	Client420_f	sad	16	0.346744309	0.51396648
Relationships	Client420_g	sad	29	0.345278726	0.627659574
Relationships	Client420_h	Friendly	26	0.349182764	0.697368421
Culture	Client420_i	peaceful	32	0.277476255	0.628571429
Relationships	Client420_j	Friendly	24	0.274651811	0.457711443
Relationships	Client420_k	admire	18	0.234877384	0.383561644
Ability	Client420_l	peaceful	18	0.242622951	0.282828283
Relationships	Client420_m	furious	26	0.270691995	0.411764706
Relationships	Client420_n	Friendly	24	0.334394904	0.741176471
Relationships	Client420_o	Friendly	23	0.328298087	0.681818182
Behaviour	Client420_p	Friendly	30	0.343945069	0.68852459
Relationships	Client420_q	peaceful	20	0.33661865	0.611111111
Behaviour	Client421_a	peaceful	20	0.423622047	0.6
Relationships	Client421_b	Friendly	23	0.379807692	0.607929515
Ability	Client421_c	peaceful	28	0.373980249	0.642487047
Relationships	Client421_d	peaceful	22	0.399896801	0.587628866
Relationships	Client421_e	peaceful	22	0.343369635	0.532467532
Relationships	Client421_f	peaceful	21	0.344683808	0.592592593
Relationships	Client421_g	peaceful	31	0.400102459	0.763636364
Ability	Client421_h	sad	29	0.380582524	0.8
Relationships	Client421_i	peaceful	29	0.389105058	0.708333333
Relationships	Client421_j	peaceful	25	0.465486726	0.709677419
Relationships	Client421_k	peaceful	32	0.378965922	0.716666667

Topic	Session_name	Emotion	Emo_value	Richness	POS
Personality	Client421_l	peaceful	20	0.381278539	0.579545455
Relationships	Client421_m	peaceful	26	0.404569253	0.615384615
Ability	Client422_a	peaceful	56	0.371589085	0.485981308
Ability	Client422_b	peaceful	17	0.335490831	0.428571429
Personality	Client422_c	peaceful	7	0.427659574	0.364705882
Ability	Client422_d	sad	23	0.362711864	0.302325581
Behaviour	Client423_a	peaceful	17	0.372093023	0.48
Behaviour	Client423_b	peaceful	21	0.278846154	0.218934911
Development	Client423_c	admire	14	0.358938547	0.363157895
Ability	Client423_d	sad	15	0.279401283	0.226950355
Development	Client423_e	peaceful	32	0.324293785	0.301204819
Ability	Client423_f	sad	17	0.296654275	0.263157895
Ability	Client423_g	sad	22	0.292759706	0.313653137
Relationships	Client423_h	peaceful	32	0.282323548	0.352331606
Behaviour	Client423_i	sad	20	0.294806145	0.303191489
Ability	Client425_a	peaceful	15	0.317727512	0.319444444
Ability	Client425_b	sad	41	0.326794472	0.47826087
Ability	Client425_c	Friendly	17	0.334569733	0.282352941
Ability	Client425_d	Friendly	81	0.318044659	0.382488479
Ability	Client425_e	sad	24	0.346594005	0.419230769
Ability	Client425_f	admire	18	0.315300546	0.275449102
Ability	Client426_a	Friendly	41	0.24757953	0.296
Behaviour	Client426_b	Friendly	16	0.299270073	0.262357414
Behaviour	Client426_c	Friendly	15	0.466192171	0.571428571
Ability	Client426_d	sad	16	0.284274194	0.330935252
Ability	Client426_e	Friendly	17	0.288987435	0.23255814
Ability	Client426_f	furious	3	0.467836257	0.297297297
Ability	Client426_g	Friendly	29	0.255057168	0.308370044
Behaviour	Client426_h	Friendly	25	0.24160632	0.321428571
Ability	Client426_i	Friendly	29	0.299909666	0.391566265
Ability	Client426_j	Friendly	21	0.218287526	0.373540856
Ability	Client426_k	Friendly	17	0.319103522	0.455445545

Table C.1: Complete results of emotion extraction and writing style features for client's sessions

Appendix D

Ethical Approval Form

Central University Research Ethics Committees

16 April 2020

Dear Dr Grasso

I am pleased to inform you that your application for research ethics approval has been approved. Application details and conditions of approval can be found below. Appendix A contains a list of documents approved by the Committee.

Application Details

Reference:	7668
Project Title:	Evaluating a conceptual framework for monitoring student wellbeing in VLEs
Principal Investigator/Supervisor:	Dr Floriana Grasso
Co-Investigator(s):	Mrs Lubna Alharbi, Mr Phil Jimmieson, Dr Floriana Grasso
Lead Student Investigator:	-
Department:	Computer Science
Approval Date:	16/04/2020
Approval Expiry Date:	Five years from the approval date listed above

The application was **APPROVED** subject to the following conditions:

Conditions of approval

Please note: this approval is subject to the restrictions laid out in the Policy on research involving human participants in response to COVID-19. Therefore all face-to-face contact with human participants for the purpose of research should be halted until further notice; unless the study qualifies as one of the exceptions specified in the Policy and has been discussed with Research Ethics and Integrity team.

- All serious adverse events must be reported to the Committee (ethics@liverpool.ac.uk) in accordance with the procedure for reporting adverse events.
- If you wish to extend the duration of the study beyond the research ethics approval expiry date listed above, a new application should be submitted.
- If you wish to make an amendment to the study, please create and submit an amendment form using the research ethics system.
- If the named Principal Investigator or Supervisor changes, or leaves the employment of the University during the course of this approval, the approval will lapse. Therefore it will be necessary to create and submit an amendment form within the research ethics system.
- It is the responsibility of the Principal Investigator/Supervisor to inform all the investigators of the terms of the approval.

Kind regards,

Central University Research Ethics Committees

ethics@liverpool.ac.uk

794-8290

Appendix - Approved Documents

(Relevant only to amendments involving changes to the study documentation)

The final document set reviewed and approved by the committee is listed below:

Document Type	File Name	Date	Version
Participant Information Sheet	Information Sheet Focus Group		
Participant Consent Form	Consent Form Focus group	10/04/2020	1
Interview Schedule	Focus Group Schedule	14/04/2020	2

Appendix E

Participant Consent Form

#4

Participant consent form

Version number & date: 1, 10 April 2020
Research ethics approval number:
Title of the research project: **Evaluating a conceptual framework for monitoring student wellbeing in VLEs**
Name of researcher(s): Floriana Grasso, Lubna Alharbi, Phil Jimmieson

Please initial box

1. I confirm that I have read and have understood the information sheet dated 10 April 2020 for the above study, or it has been read to me. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
2. I understand that taking part in the study involves the participation in a video recorded focus group	☐
3. I understand that my participation is voluntary and that I am free to stop taking part and can withdraw from the study at any time without giving any reason and without my rights being affected. In addition, I understand that I am free to decline to answer any particular question or questions.	☐
4. I understand and agree that my participation will be video recorded and I am aware of and consent to your use of these recordings for the following purposes: production of the transcripts.	☐
5. I understand that I can ask for access to the information I provide and I can request the destruction of that information if I wish at any time prior to final submission of the PhD dissertation. I understand that following this I will no longer be able to request access to or withdrawal of the information I provide.	☐
6. I understand that the information I provide will be held securely and in line with data protection requirements at the University of Liverpool until it is fully anonymised and then deposited in the University of Liverpool Library as Appendix to the PhD Dissertation for sharing and use by other authorised researchers to support other research in the future.	☐
7. I understand that my responses will be kept strictly confidential. I give permission for members of the research team to have access to my fully anonymised responses. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report or reports that result from the research.	☐
8. I understand that confidentiality and anonymity will be maintained and it will not be possible to identify me in any publications.	☐
9. I understand that signed consent forms and original video recordings will be retained in the University Sharepoint site until the deposit of the final version of the Dissertation.	☐
10. I agree to take part in the above study.	☐

#4 UNIVERSITY OF LIVERPOOL

___________________________	__________	_______________________
Participant name	Date	Signature
___________________________	__________	_______________________
Name of person taking consent	Date	Signature

Principal Investigator
Floriana Grasso
Department of Computer Science
University of Liverpool
floriana@liverpool.ac.uk

Student Investigator
Lubna Alharbi
Department of Computer Science
University of Liverpool
lubna@liverpool.ac.uk

Appendix F

Participant Information Sheet

Evaluating a conceptual framework for monitoring student wellbeing in VLEs

Version 1. 10 April 2020

You are being invited to participate in a research study. Before you decide whether to participate, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and feel free to ask us if you would like more information or if there is anything that you do not understand. Please also feel free to discuss this with your friends, relatives and GP if you wish. We would like to stress that you do not have to accept this invitation and should only agree to take part if you want to. Thank you for reading this.

1. What is the purpose of the study?

The purpose of the study is to informally evaluate a conceptual framework for monitoring online students' wellbeing. The framework has been devised by a PhD student of the University of Liverpool, and is only theoretical, has not been implemented. The study aims at finding out whether practitioners in the field would find the idea appealing and feasible.

2. Why have I been chosen to take part?

You have been selected to take part because of your recognised expertise in online learning.

3. Do I have to take part?

Participation is voluntary and you are free to withdraw your participation at any time, without explanation, and without incurring a disadvantage.

4. What will happen if I take part?

You will be invited to participate in a mini-focus group study, alongside other two participants, and moderated by one of the researchers in the project. The focus group will be online, via MS Teams, and the session will be recorded, to aid the production of transcripts. During the focus group you will be asked to provide some feedback on the theoretical framework, by drawing on your practice and expertise.

5. How will my data be used?

The University processes personal data as part of its research and teaching activities in accordance with the lawful basis of 'public task', and in accordance with the University's purpose of "advancing education, learning and research for the public benefit.

Under UK data protection legislation, the University acts as the Data Controller for personal data collected as part of the University's research. The Principal Investigator acts as the Data

Processor for this study, and any queries relating to the handling of your personal data can be sent to Dr. Floriana Grasso floriana@liverpool.ac.uk.

Further information on how your data will be used can be found in the table below

How will my data be collected?	The focus group session will be recorded using the function on MS Teams.
How will my data be stored?	The recording, and the transcripts subsequently extracted, will be kept on the University Sharepoint site.
How long will my data be stored for?	The video will be destroyed as soon as the final version of the PhD dissertation will be deposited. The transcripts will form part of the Appendix to the dissertation.
What measures are in place to protect the security and confidentiality of my data?	The transcripts will be anonymised. The video will not be made public and will be destroyed.
Will my data be anonymised?	The transcripts will be anonymised.
How will my data be used?	Quotations from the focus groups will be used, after being anonymised, as a form of evaluation in the PhD dissertation. They may also be used for any publication deriving from the dissertation.
Who will have access to my data?	The video will be accessible to the three researchers on the project solely.
Will my data be archived for use in other research projects in the future?	No.
How will my data be destroyed?	The video will be permanently deleted from the Sharepoint site.

6. Expenses and / or payments

You will need internet connection to participate to the focus group. We are not able to reimburse any costs associated with your participation.

7. Are there any risks in taking part?

There are no risks in taking part to the focus group, but should you experience any discomfort or disadvantage this should be made known to the researchers immediately.

8. Are there any benefits in taking part?

There are no intended benefits to the participants to the focus group.

9. What will happen to the results of the study?

Quotations and paraphrases of the comments in the focus group will be used in the Evaluation chapter of the PhD Dissertation, and may be used in publication deriving from the dissertation. The Video will be destroyed. You will not be identifiable from the quotations.

10. What will happen if I want to stop taking part?

You can withdraw their participation in the study at any time, without explanation, by exiting the MS Teams session. The recording up to the period of withdrawal will be used, unless request that no further use is made of it: in this case, only the comments of the other participants will be used in the transcripts. In order to request this, please contact Dr. Floriana Grasso floriana@liverpool.ac.uk.

11. What if I am unhappy or if there is a problem?

If you are unhappy, or if there is a problem, please feel free to let us know by contacting Dr. Floriana Grasso floriana@liverpool.ac.uk and we will try to help. If you remain unhappy or have a complaint which you feel you cannot come to us with then you should contact the Research Ethics and Integrity Office at ethics@liv.ac.uk. When contacting the Research Ethics and Integrity Office, please provide details of the name or description of the study (so that it can be identified), the researcher(s) involved, and the details of the complaint you wish to make.

The University strives to maintain the highest standards of rigour in the processing of your data. However, if you have any concerns about the way in which the University processes your personal data, it is important that you are aware of your right to lodge a complaint with the Information Commissioner's Office by calling 0303 123 1113.

12. Who can I contact if I have further questions?

Dr. Floriana Grasso
Department of Computer Science
University of Liverpool
Tel. +44 (0)151 7954240
floriana@liverpool.ac.uk

Appendix G

Interview Schedule

UNIVERSITY OF LIVERPOOL

Evaluating a conceptual framework for monitoring student wellbeing in VLEs

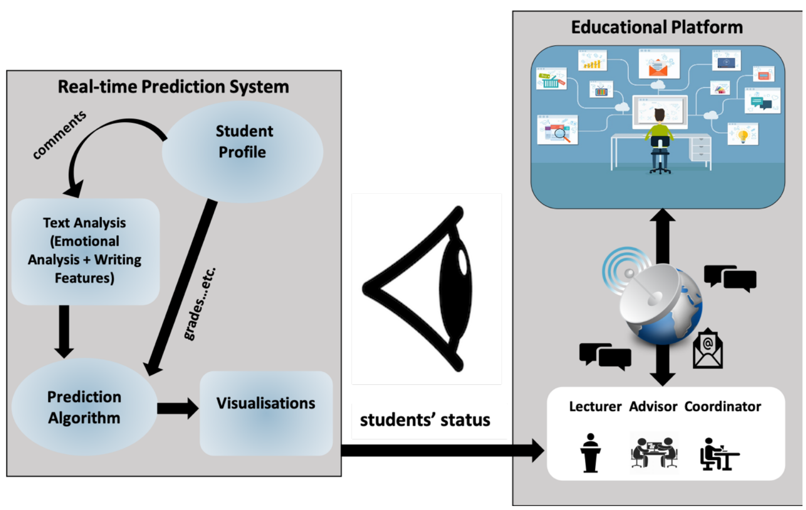
Version 2. 14 April 2020

Focus Group Question schedule

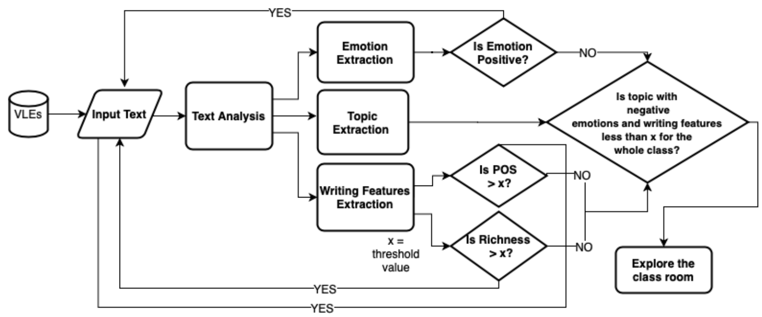
Note: None of the questions elicited any discussion of specific cases of student well-being or were in danger of identifying people - these were strictly hypothetical discussions about the proposed system/software.

Introduction to the purpose of the research.

1. Describe your expertise: why do you think you can contribute to this discussion?
2. Online students and mental health: have you perceived this as a problem, more so recently, and have you noticed a deterioration, or a different frequency of reporting?
3. What is your experience of using learning analytics to feel the pulse of a classroom, or to gain an understanding of how a particular student is progressing?
   - follow up: for online students, on your experience, does the analysis of their online behaviour give a good ground to make a judgement on any personal or learning challenges they may experience?
4. Have a look at this picture: what is your immediate reaction to it, and why?



5. Have a look now at this flowchart: what is your immediate reaction to it, and why?



6. What would be the challenges or the barriers to the update of a system like the one depicted in the images you have seen?
7. If you could have access to an ideal system, what would be the single feature you think the system should include?
8. Do you have any other comments, or any issues you want to mention which has not been mentioned so far?

Thanks for your participation.

Appendix H

Responses to the FGD Questions

1. **Question 1**

   - **Participant A:** I have been involved in online teaching and learning since 1995. I started working in the commercial area in online format and then eventually moved to the academic world in 2002. I am currently working as a lecturer in Higher Education and innovative use of technology in education. I have been involved in the overseeing [*University name omitted*] online programs as a partnership with first [*company name omitted*] and then also [*company name omitted*] when they took over, I worked on master's degree program and computing(the EDV program), so I looked at online teaching and learning from two different disciplines perspective. I also worked in online learning for her own campus environment as well as working in one of the other partnerships, the [*University name omitted*]. Then, in another commercial, for profit institution in the state of [*omitted*] as well. I had a kind of interesting background in terms of working for different types of institutions. I have been working in the commercial IT area, first for Microsoft and then I continued to do research, conduct research on the notion of online presence for faculty and students.
   - **Participant B:** I had a number of years working on campus with students, with some teaching, but mainly in a support role. Then, I worked for 14 years in online master student primarily in student support and leading teams of advisor who are available for students studying online. I could offer some insight to how best able to do that how they support the students and what he saw on the campus worlds. Also, I have spent some time looking at predictive analytics and that's an area of increased research.
   - **Participant C:** I worked for a university face to face and retired in 2013, and during that time I were doing some part time online work for the [*University*

name omitted] until 2013. Then, I became full time as the program director for the director of online studies at the [*University name omitted*]l. And in 2019, I moved to [*University name omitted*] as a full time program director of their master's programs and their doctoral program. So, I have had some significant on ground teaching and about 20 years worth of online program.

2. **Question 2**

- **Participant A:** I had students who have come to me with mental health issues, and that's a remarkable scenario given the fact that we are working in a very multicultural environment online at this particular program, because so often around the world to discuss mental health issues is culturally not considered to be acceptable.

  The fact that students are coming from different parts of the world and feeling. Are they feeling confident enough or comfortable enough? It is another word that I would use where they have expressed a concern that they have had. And often the conversation is driven by the lack of sufficiently promising progress in their academic work.

  And so they are coming with originally sometimes they comment saying, "look, I am really sorry, I have not done a very good job.". That would be the kind of thing they would say my performance is not good.

  But then what they would do is they would then coach couch some of it in terms of giving me an explanation that could or could be interpreted as an excuse. But then as they unpack it even more with me, often when you are reading what they are saying or sometimes they actually want to talk with you because, you know, I do talk with students.

Sometimes they will get into a little more detail. Then, quite frankly, I am a little surprised that they are. But they do they do come in. And then often what they talk about is they talk about the stress of trying to manage their family or trying to manage their work. So often when they come to me with these kinds of concerns, they are driven by performance. But when they start to unpack it with me, it it does not always work. Although you would think that would be work driven, it is not because something has happened in their personal life.

I have had a student come to me and explain in some detail. Interestingly enough, "a man" So when I think about the gender stuff, a man is actually coming and telling me about their divorce and they are coming from a culture that does not welcome a conversation about divorce if they are male. Also, I have had a female come to me with the same topic.

I have had students come to me about family concerns or mental health issues, and then I have had students actually come to me and say, I have to take a break. I have a mental health document that I will share with you, but I ask for privacy, things like that.

So, have I seen an increase? I am not sure if I can say an increase, but I have seen as an increase of being acceptable to talk about it in a way that maybe they did not in the past. Maybe that is because we have technology that allows them to feel more connected now because they can meet beyond just the text. I definitely have had students come and talk to me about mental health issues.

- **Participant B:** It sounds really familiar for me. It would be struggle to say whether I have noticed that it is got worse, or if you have seen an increase in the volume of students that are having difficulties, primarily because certainly with the online work. The time that I have worked those 14 years and I started working with a relatively small amount of students and at one point it increased from certainly ten fold. So then the number of cases that you come across is obviously more. But whether it was because the situation is getting worse and more students were suffering mental health issues, I am not sure.

But I do feel that there was I would agree with Participant A. There was certainly a lot of cases or a number of cases where students would come forward with talk quite candidly about their experience, what they were going through. I also wonder sometimes if it allowed it, if it made it slightly easier. I think sometimes that coming into the office of a lecturer on campus, sometimes it might just come out as a third of an emotional outpouring feeling because of the situation, whereas I think online allows much more measured perhaps. I think some people were quiet, really quite open, but would present themselves in a slightly more open way, particularly and culturally often that is it is not acceptable.

To your point, Participant A, but feel a little bit safer in the same way that we talk about cyber bullies, that maybe it is getting worse because it is so much safer somehow easier to do it behind a keyboard or a phone than it is face to face or difficult several. I saw an increase close on a lot of students forced me to come forward and talk about some of the issues they were facing.

- **Participant C:** There is a question of what do we mean by mental health? Because mental health be a lot of things. I maybe see the one online are less stressed than the ones in class because in class they are doing now right away, they are not used to it.

  But the online, their stress is probably more related with what is happening in the home as a opposed or sickness.

  I will say that having been at [*University name omitted*] for so long and then now being at [*University name omitted*] I do see some differences in how they have address "the mental health", the issues that students have about being present in the classroom and and getting their assignments set. because

In [*University name omitted*], it is quite rigid, if the student can not meet deadlines because of health issues, they have to go and do a petition. They have to do an exception and then have to wait for a days, maybe a month or two to really find out what has happened as a result of whatever was happening in the classroom. Years ago when I were at [*University name omitted*], I remember a student was from Ghana where they had an uprising and their health was burnout and at that time, we said, OK, take the week off and then get caught up when you can. We let them have that flexibility.

But We do not have that now, We have not had it for a few years at [*University name omitted*]. And as a result, a lot of the mental health, I think is exacerbated because they can not find out what is going to happen to them for a month or two or three.

In [*University name omitted*], they already have very immediate petition's experiences so that within a week their petition is looked at and reviewed by a team and then agree to or not accept their excuse. And in many of the case there, obviously, if they have medical documentation and other things that are associated with their mental health, they know right away. And I think that has an impact on how under mental health. At least I am seeing it from a comparison of the two.

- **Participant B:** I think it is an interesting point Participant C. You mentioned something that I and Participant A that I think we had. I worked with a lot of students who, when they were embarking on their studies might disclose some particular issues that they face chronic "health issues, mental health issues".

But what was interesting is if even that was the case, if they declared something or they did not, the impact that then certain things that can happen either in their world, in their private world, personal world or in the online study world exacerbates everything so greatly. I think that really are the moment for whatever reason, "the assignment is not clear" so they have to spend a little bit more time working out what it is, or "the library service down.".

Then suddenly, I think a lot of the times the experience I had students would talk about some of the perhaps mental health issues that they face. They would come up when those kind of additional factors happened. So, you would hear "I am already at the level with my work and my health and now I have been trying to work, I have to log into the library for the last two hours, you can not do this to me, this makes me ...". and you hear and see the anxiety, the outpourings, long emails or distressed phone calls.

So that's balance between an online world where at the mercy of the technology or the that was where I felt the thought that sometimes it would concern all their underlying health issues that have not been disclosed, because sometimes the reactions to relatively small issues or delays that can happen online world became these last things. The balance of play between individuals health and then health going in the course of a classroom is a really important point to make. It could have all kinds of consequences.

And we might say, well, the server was down for 30 minutes. Deal with it. When somebody is up against a deadline and juggling all kinds of other things, competing with it, it is fellow than that is quite often that's when students would come and say, well, I have these issues and you can not do this to me because I am already struggling. So you have to get everything right for me not to explode or more imploded of these.

- **Participant A:** I am hearing you saying that it may also be contextual in terms of the online environment. Add in the specific structures, add in the locations and the students around the world. Add those all into the online context that we are talking about specifically referring to. May be all factors that contributed to increased worry or stress because we are actually engaged in a somewhat more stressful environment to begin with.

  I mentioned it all and because last night the server went down and we worked out as it is always done. Before we workout I go around and I emailed all of my students to say, "here is the workaround" and I got things back saying, "thank you, Participant A, you are a lifesaver,I was been trying for four hours now."

  Why was this such a big deal? If it been Friday, they probably would not even notice. Why was it a big deal yesterday? Sunday night submission. So I am wondering, one of the things that me we are talking about is that particular context in which we are working somewhere where it may be creating stressful and increased levels of feeling of mental health concerns, where if you transplanted them into a slightly different environment such as MOOCs, where it is a pretty free form kind of environment, the mental health concerns might be different.

3. **Question 3**

- **Participant B:** I think it is a good question, as certainly I have seen cases that I think one of the things that you can absolutely do to try and support online students as much as possible is to look and to have people that are working closely with our students to be aware of it and how do you help them identify changes in behaviour, changes in tones, uses of language that indicates that there might be something a miss? I think you can absolutely do.

And I have certainly seen cases of it. I think one of the things that is absolutely key is of course, it is sentence because it is recorded. So I had a case, two years ago, where a student wrote to us to tell us that he was in a very, very bad way, was going to commit harm to himself.

It was very distressing for the support person when you have someone, it is so clear in writing, it is a very extreme example of that. We have more insight. We get to know students more through their online behaviour and their online work because everything is visible. And we had one or two cases where support people would receive messages and they would be able to go to their manager or to colleagues. I have received this messages from student, here is some of the previous correspondence from this student. I have some concerns because of this.

And then you are able to just take a few minutes, have one or two people to assess it just to see how best to approach it. Who is it now? What contact do we need to make the student revolt? I think that is a big advantage of the online world compared to on campus, face to face from one person can suddenly be in a situation that may be unsure as to how to deal with it.

We certainly have the opportunity where you could take your time to assess behaviour change. Are we seeing something that means the students perhaps getting into trouble? So we certainly saw I think from a faculty point of view, you probably saw it with so many interactions in class.

- **Participant C:** I think often it settle, but you do get the feeling that a student is not in a good place. The idea that maybe and it may even come through a discussion questions where they will share something very personal and you think, oh, you know, this is going to make everybody uncomfortable. But you can see behind that what they have written that they are in a very bad place and they really need to step back and talk to somebody about it.

But from an instructor point of view, we are kind of limited about what we can do with these students or in a class. I mean, I think that there have been some times when I contacted student support and said, hey, you should contact that student. Because there is something going on with him or her. But but the instructor themselves, probably are not the contact person if they are going through mental health, unless they are the dissertation adviser. And there you know that student really well, and if they are going through something, you know they are going through it.

- **Participant A:** Although I will say that I would agree with Participant C, but I would also say I have had experiences where I have picked up something, a mess that is not typical as a student and often it is near the end of the module. So I have had some level of engagement already. I have made a comment on my feedback. They have written me back and saying, how did you know? Here is what is been going on.

  And it was not specific my comment. It was more generic that it was not going to be offensive in case there really was something going on. It is just a comment, just as a personal one to one kind of comment. We do expect faculty to do it all. I think that is probably a reaction or a response, because some of us are more in tune with our students.

  I have had situations where students have written me back and said, how did you not, like in a private message. How did you know this is what is really going on? I am so glad you noticed it or whatever. It is like the door opening. So on campus, often when a student stands in your door and really wants to talk, even though it is not office hours, you know that that is the time you let him in because, you know, you never know really what they are going to ask about. So I do. I have had that situation. But so often it is because either I have noticed it, it is out of character. I think it is probably a good way to say it.

- **Follow-up Question(3):**

  * **Participant C:** Well, I think, like I said, it is kind of stand up, if you have been doing it for a while and you have been tuned in to the student. You know, when something is different. It is kind of like when, you know, when somebody submits the paper or submit to a discussion question, you are absolutely know they did not do it. You can not say it is because of this. You just know that student would not have done it. It is kind of that way when you are looking at students in term of maybe their online personality. When I taught the first class, I always said you all have your own personal online identity and personality.

    And so if something seems off on that personality. Like Participant A said, there are not everyone, not every instructor would be looking for that or seeing that. But I think that the ones that have been around for a while. But essentially we do get the feeling of who this person is. I do. I mean, in every class I taught, I kind of had , whether they would like that in person, I don't know. But their online personality was pretty clear. And so when it when they are not in, when you do not see that same person, then something is going on.

  * **Participant A:** Often it is how they are write things. Their style of writing. How they phrase, how they present an idea. If their typical

approach is to be logical. And then they stop writing it in logical format that you are used to seeing. It is not always you know, it is often it is a tip that they are under stress of some sort.

4. **Question 4**

- **Participant B:** I think the idea that we can potentially know the students online through all of their footprints "the millions and millions of footprints" left by every single click. All of that online behaviour allows you to build in those predictive analytics to be able to incredibly efficiently, accurately identify what the particular behaviours.

  The outcomes might be occur because of the result of what you see online. That is my thought. I was going to say before we reach started about what you are all talking about in terms of the online persona. That is what we have always worked over the last 10-15 years on having as much as possible students supported by the same person throughout their journey, so that you have got to build that relationship. You have got to understand that person quite well. Then, the support person would have regular contact. We reach to individual students, get to know them. They would be best placed to identify anything that might be seen as a change in behaviour or a sign of concern, but also that it builds that relationship for students to open up to talk them about it.

  If there was something that was troubling within this, I felt like I built quite safe relationship. This is the next phase for me, says everything about how once you can do that very well with people who are trained and then those may be good in a support role.

  The next level of me is to use and utilising the technology. Using of all online footprints that students leave behind to be able to help those support people and help those responsible for the support and the welfare of the students to identify things change. So that is what this looks like to me. I think it could be absolutely critical future of helping identify those students that might need some assistance.

- **Participant A:**

  I have understood right now looking at learning analytics as a vehicle for real time interventions within the learning platform. So, they are not necessarily looking upfront at the real time at the predictive piece from a holistic perspective, but only looking at what the student does once they are in the learning platform with the same idea that you could then produce a much more personalised intervention to a student while they are in the middle of learning as supposed to understanding the student on an ongoing basis, having a profile at the start and then predicting how they might perform down the road.

  So when I am looking at this, you are gaining a sense of the student first, perhaps from historical perspective, but also what they are currently engaged in and then being able to feed that status on a continual basis to the educational environment. So I see the platforms spin stripped out slightly differently in terms of there is a facilitator role of some sort (lecturer, advisor, coordinator).

  Then there is the actual platform itself that all that thing can be part and parcel of a potential world of engagement. I guess is probably a good way to say, you know, I have a certain engage within their learning environment based on what we know about them. That is kind of in addition to what Participant B was saying, that is kind of what I am seeing here.

- **Participant C:** I am not seeing anything here like that. I am not sure what I am saying here, but if I am going to be honest because if we have a real time prediction system. And we have a student profile which includes whatever they have done in classes. And then we also have text analysis. I am not exactly sure what we are doing with text analysis. Is this an external person or who is getting this information?

There is obviously somebody must be getting this information to go into the prediction algorithm. And then all this is going to go to the lecturer. I do not think so. The average lecturer does not see anything about a student as if we were going.

If we in the real sense we are not an instructor may have never had this student before. Especially, if we have a lot of students. So the only thing the instructor sees is what that student brings to that classroom today for this week, or eight week term. So, I am not sure what we are doing here with this prediction system for the lecture.

- **Participant B:** So I saw it as what we could or might not necessarily be doing in an online platforms right now, but the work could help. That is why we are saying that what I saw is really important to the support staff for building one to one relationship and getting to know the personalities and getting to potentially prevent or see signals of future difficulties.

  But this would be, we have begun with predictive analytics, mainly around just grade performance. So our engagement in the classroom. I see this as something that you could expand to things like health issues or mental health issues. But then you could we are my in my role anyway. We decided not to make it available to, for instance, faculty in class. But I see this as anybody that is interacting with that student on a regular basis who could be benefit from the warning signs in the prediction, I guess.

- **Participant C:** Well, I do think that this will be useful predicted. But, I am just not sure the lecture would get any of this information about the students status. Or even if we want them to.

- **Participant A:** Well, I guess so. So what I am moved out right now is the current paradigm that I am working in, and I moved into the paradigm where I am seeing this being used actively through my side, through the eyes of my doctoral student.

  And one of the things that you are proposing is that if indeed your research is turning out what you think is doing through regression analysis. So you are looking at a bunch of retrospective data you are using because there is ethical issues, by the way. This is fraught with all kinds of ethical issues as well. They are related to categorising or pre-judging students.

  But one of the things that you are proposing is the system that you are pulling this data from is actually allowing faculty to use the data in real time. But the educational platform has been built to allow them to do that. So if I take myself out of the current platform and put myself into what if we had an educational platform that allowed real-time knowledge about students and shifts in students behaviour through the use of analytics, what potentially could that do in terms of improving in our ability to intervene in a more timely basis or potentially in a slightly different basis than what we would have ourselves even thought about it

- **Participant B:** I think that is a good point, because I was just thinking, looking at it, thinking how far those ethical issues that you talk about, in terms are easy to think about anything that you can do additional do to help support students and help prevent issues. But I imagine all of the potential pitfalls and dangers of living in a world where online personality can be relatively accurate and maybe not entirely accurately predicted as all.

This student is going to or is suffering from central mental health issues and how you approach that when maybe it is so settle then an algorithm is picked up that issues are on their way, but the individual themselves would be so far away from me knowing that what you do with that information.

I can see so many benefits about the support that you could put into place. But I think it is really important issue around. I do not think students signing up for a study at online university or possibly even necessarily aware how much data you can gather and able to gather from them to build up your model of who they are and what their behaviour might be.

It is some super strong points, ethically very challenging term to get the balance. I think between support and invasion of. All kinds of privacy.

- **Participant C:** Okay. if we change the word student to faculty. You could see this working really well for a program director. Because Participant A was a faculty manager, we did know when a faculty member might be dealing with something because something was going on in his classroom. And it was clear. I mean, it was like, okay, something is happening with this faculty member.

  He is not doing the same things he is done in the past. He is not here. The way he writes is different. The way he interacts with the students are different. It was a pretty clear to me when a faculty member was either getting burnout for some reason or they had home issues. Whether that were affecting their mental health, you know, maybe their wife was pregnant and she miscarried or maybe they had a car accident and it was affecting their mental health as well as their their physical health.

  So this could be more predictive for a faculty member. And we do not have that many ethical issues. Well, we have definitely privacy issues, but not when you are there supervised. When you are supervised, when you are their supervisor, you have access to all of this, because if they are having issues in a classroom, you can see everything they have written and you can tell you can do a text analysis to see if it is different. If you do not know if we want it. Participant A, you can say something about this, but I mean, I am not sure I could put it into a numeric value, but the difficult value. But I can definitely see the difference.

- **Participant A:** The one issue that came up consistently was issues around burnout. And we early on the programs we would lost couple of really top notch faculty because of burnout. And they just left. I mean, that is how bad it was. And one of the things then that in my role as faculty director, faculty involvement person for a while was keeping an eye on burnout because, we often talk about student burnout. But there was very much cases of faculty burnout.

  We had a number of little tricks that we used to do in terms of how often did they log in? What were their postings like? How did they use the technology? How did they use the words in the structures of the sentences? How many students were they managing? What classes? And what is happening? So, in other words, what was the class and then how many students.

  So I used to keep a running total of faculty and what classes they taught. And I still do kind of that record. I still do in it now because I am doing scheduling and I do it because I learned very early on in the program back in 2004-2005 that faculty burnout. And so there were ways to catch them before they got themselves into a point where they burnout rather than losing. I wanted to make sure they did not get burnout.

  So with this kind of system, help anybody who is trying to keep track of how well faculty are doing while they are working alone, I think the answer is yes. I think the variables be slightly different, but I think certainly would be something that be helpful.

5. **Question 5**

- **Participant C:** I think this would be probably more difficult with the [*University name omitted*] model because of the multicultural aspect of their students. I mean, you will find that some students are very emotional about something sometimes it is a little off the wall. But it still pops up like, say, for example, they are a true moralist because of their culture. So everything is black and white, not grey. And so they write that way. They write about whatever topic they are talking about in a very black and white way. And that may come across as being non emotional or very emotional, one or the other. It might flag that like if they are in there but it may not have any. It may not really be anything. It is just the way they their culture.

- **Participant B:** What I imagined something like the part in the predictive algorithm, based on the previous graphic that you showed, that was a text analysis that you would have algorithms working in in one particular area to identify certain things that might be a trigger a flag to then explore what was actually going on? So from my experience of I have looked at predictive analytics from big data to try and predict student performance.

  Then one of the things, one of the significant bills, an enormous regression model and one of the significant factors that was a successful predictive factor was a language analysis, automatic language analysis on the complexity of grammar and use of language and better somebody schooled in that. It was a better predictor of succeeding in an online classroom.

So this piece felt to me a bit like that was part of what was alluded to in the previous graphic, that that would be the kind of emotional. And so it feels to me like if that thing is going to be possible and could be super beneficial to put less pressure on those individuals who are supporting students to be looking for those often very, very settle, like I said, very settle differences in behaviour or language and emotive language in particular, to be able to say.

I think actually through all of the data driven lies or this individual, you should go explore the classroom because I saw this as a part of that on the left, on the site and here the graphic ready. So it feels to me like something you would have to be done automatically through complex language analysis. Well, I was thinking that it would be able to be tested in testing so that you could be a certain as you could be that those flags would be as accurate as possible and still need human follow up.

- **Participant C:** What would be more effective is to run the analysis against a particular student. So, for example, a student within X class and then they were in X X class and then there were an X X X class. And then you look at there analysis of their text across those three courses and you will see if there is any issues. That popped up.

- **Participant B:** From looking to predictive analytics, we ran for student performance twice and I hate to ask the question, how scary is this? Because as a result, it was scary enough to just look at predicting student performance.

  And it is scary. I mean, just that the magnitude of it. And then the accuracy with which you could do it. So, yeah, to your point, Participant C, and we could choose at any point, however many times a day we wanted to just import all of the data that students are leaving behind in our classrooms. And at any point we could score. How likely that student was to get into difficulties. Now, that was to do with grading performance.

And so we would just as a process student, although students from when we decided we wanted to run those analytics and when we saw particular flags, we would contact them to try to head off some of those difficulties. And when you talk about what that would mean for potentially around mental health and those behaviours and the kind of emotional things you analysing it, it really is.

I mean, it would be fascinating, but it is getting an insight that I think about how we as emotional beings struggle sometimes to understand our own emotions or be able to talk to him and be able to open up this kind of level of analysis where any point, somebody can run the analysis and say, yep, there is a red flag. And finally the last weeks, but he is about to explode. Somebody got to it quick. It is relatively scary, but also potentially super, super positive. it is just constantly going on, that you could constantly run it to go. It is not like you decide to go and analyse one student president of a school for every single person in your mind as to how they were doing.

6. **Question 6**

- **Participant A:** Well I think you would have to look at the factors. The first fact you would have to look at is if you are going to be so heavily dependent on the text, because that is one of the downsides to working with in an online environment that is heavily oriented towards text is that you would have to factor in issues around the English as a second language. We have I mean, we have faculty who got themselves into trouble performance wise because their use of English was not problematic in terms of it was just poor work romantically. But the rearrangements of the words were such that they provoked problems because unintentionally provoke problems because they just did not understand that when you switched words around in particular order, it went from being a positive reinforcement to a negative problem. So I think you would have an issue around language you start. I think the other one, of course, is you would have to have a factor in there, some sort of cultural norms.

7. **Question 7**

- **Participant B:** It is a great question. And I am just gonna draw on my experience of how we essentially made sense of enormous amounts of data. Really simple that if you are focusing on this one issue, you want essentially a single score as much as possible for any particular student.

  Then you can decide when you have condensed enormous amount of data that essentially identified a risk score for a particular student, then you could understand which particular groups you want to look at in more detail.

  But I think unless you go really simple, you will not be able to comprehend that you will not be able to do anything with it. If you bring it down to one simple score that then you decide on. Just based on that score, you can have a look at students, then it is possible to say this one percent of students we are going to have a look at and exploring it in a bit more detail.

  And you can look, what is it? Which variable is it? That is caused that score to go high? Is it because of something is come out of the emotion extraction? Is it because it is something is come out of the writing feature extract and you can understand that bit more before you go explore.

  But for me, I would want a single score that would enable me to go. Who is it that need to look at any more of the levels? And I thought as soon as it is multiple, the ability to decide who you want to try in to look at would be impossible. So absolutely a single score for me. Thank you.

- **Participant C:** I like that idea. I think I would also like to know trends. I do not know how you show trends, I do not know what kind of trend. Let's say you have that one percent that has the score range that we care about. Then what has been happening with those students in that one percent? I do not know how it would be represented on a dashboard. But for me as a program director, I would like to see. What is going on with that one percent in a visual way.

- **Participant A:** I think in addition to what Participant B and Participant C have said, I think that what would be helpful for me in working with data visualisation tools. What I find is I want to be able to have sufficient levels of categorisations that I am interested in. So in other words, if I were to look at this, I would use it in terms of presence. So I use theory driven categorisations of that are indicative of the different levels of presence.

  And then I want to be able to manipulate them up and down to be able to see if there is a trend. As Participant C said, and if there are absolute markers that data visualisation tools can give me, as Participant B mentioned, which is, you know, at the end of the day, here is the one that I want to look at, but I want to be able to drill down a little bit more to see what is contributing to that one.

References

[1] Ahmed Abbasi and Hsinchun Chen. Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace. *ACM Trans. Inf. Syst.*, 26(2), April 2008. ISSN 1046-8188. doi: 10.1145/1344411.1344413. URL `https://doi.org/10.1145/1344411.1344413`.

[2] Sadia Afroz, Michael Brennan, and Rachel Greenstadt. Detecting hoaxes, frauds, and deception in writing style online. In *2012 IEEE Symposium on Security and Privacy*, pages 461–475. IEEE, 2012.

[3] Per Ahlgren, Bo Jarneving, and Ronald Rousseau. Requirements for a cocitation similarity measure, with special reference to pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, 54(6):550–560, 2003.

[4] Ane Alberdi, Asier Aztiria, and Adrian Basarab. Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *Journal of biomedical informatics*, 59:49–75, 2016.

[5] Alexander Street Publishing. Counseling and Psychotherapy Transcripts, Client Narratives, and Reference Works. `https://asp6new.alexanderstreet.com/psyc/psyc.index.map.aspx`, 2008.

[6] Lubna Alharbi, Floriana Grasso, and Phil Jimmieson. A conceptual framework for identifying emotional factors leading to student disengagement in VLEs. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE 2019, page 293, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368957. doi: 10.1145/3304221.3325570. URL https://doi.org/10.1145/3304221.3325570.

[7] Lubna Alharbi, Floriana Grasso, and Phil Jimmieson. A conceptual framework for real-time emotional-state monitoring of students in vles to identify students at risk. *Human and Artificial Intelligence for the Society of the Future*, 2020.

[8] Ebba Cecilia Ovesdotter Alm. *Affect in Text and Speech*. PhD thesis, University of Illinois at Urbana-Champaign, 2008.

[9] Nabeela Altrabsheh, Mihaela Cocea, and Sanaz Fallahkhair. Learning sentiment from students' feedback for real-time interventions in classrooms. In Abdelhamid Bouchachia, editor, *Adaptive and Intelligent Systems*, pages 40–49, Cham, 2014. Springer International Publishing. ISBN 978-3-319-11298-5.

[10] Saima Aman and Stan Szpakowicz. Identifying expressions of emotion in text. In Václav Matoušek and Pavel Mautner, editors, *Text, Speech and Dialogue*, pages 196–205, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. ISBN 978-3-540-74628-7.

[11] Kimberly E. Arnold and Matthew D. Pistilli. Course signals at purdue: Using learning analytics to increase student success. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, LAK '12, pages 267–270, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1111-3. doi: 10.1145/2330601.2330666. URL http://doi.acm.org/10.1145/2330601.2330666.

[12] Kimberly E. Arnold and Matthew D. Pistilli. Course signals at purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, LAK ,Äô12, page 267,Äì270, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450311113. doi: 10.1145/2330601.2330666. URL https://doi.org/10.1145/2330601.2330666.

[13] Pam Arroway, Glenda Morgan, Molly O'Keefe, and Ronald Yanosky. Learning analytics in higher education. *EDUCAUSE*, 2016. URL https://library.educause.edu/~{}/media/files/library/2016/2/ers1504la.pdf.

[14] Anthony R Artino Jr. Emotions in online learning environments: Introduction to the special issue. *The Internet and Higher Education*, 15, 2012.

[15] Anthony R Artino Jr and Kenneth D Jones II. Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *The Internet and Higher Education*, 15(3):170–175, 2012.

[16] John T Avella, Mansureh Kebritchi, Sandra G Nunn, and Therese Kanai. Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning*, 20(2):13–29, 2016.

[17] Harald Baayen, Hans Van Halteren, and Fiona Tweedie. Outside the cave of shadows: Using syntactic annotation to enhance authorship attribution. *Literary and Linguistic Computing*, 11(3):121–132, 1996.

[18] Bonny Barr. Identifying and addressing the mental health needs of online students in higher education. *Online Journal of Distance Learning Administration*, 17(2), 2014.

[19] Rachel Batchelor, Emma Pitman, Alex Sharpington, Melissa Stock, and Eilidh Cage.

Student perspectives on mental health support and services in the uk. *Journal of Further and Higher Education*, pages 1–15, 2019.

[20] David C Berliner. What's all the fuss about instructional time. In Miriam Ed Ben-Peretz and Rainer Ed Bromme, editors, *The nature of time in schools: Theoretical concepts, practitioner perceptions*, volume Part I, pages 3–35. Teachers College Press, 1990.

[21] Mudit Bhargava, Pulkit Mehndiratta, and Krishna Asawa. Stylometric analysis for authorship attribution on twitter. In *International Conference on Big Data Analytics*, pages 37–47. Springer, 2013.

[22] Marie Bienkowski, Mingyu Feng, Barbara Means, et al. Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *US Department of Education, Office of Educational Technology*, 1:1–57, 2012.

[23] Haji Binali, Vidyasagar Potdar, and Chen Wu. A state of the art opinion mining and its application domains. In *2009 IEEE International Conference on Industrial Technology*, pages 1–6. IEEE, 2009.

[24] Joanna Blaszczak. *Investigation into the Interaction between the Indefinites and Negation*, volume 51. Walter de Gruyter GmbH & Co KG, 2019.

[25] Avrim L Blum and Pat Langley. Selection of relevant features and examples in machine learning. *Artificial intelligence*, 97(1-2):245–271, 1997.

[26] Vasiliy Boychuk, Kirill Sukharev, Daniil Voloshin, and Vladislav Karbovskii. An exploratory sentiment and facial expressions analysis of data from photo-sharing on social media: the case of football violence. *Procedia computer science*, 80:398–406, 2016.

[27] Sandy Britain and Oleg Liber. A framework for pedagogical evaluation of virtual learning environments, 2004. URL https://hal.archives-ouvertes.fr/hal-00696234. Research report.

[28] British Educational Research Association (BERA). Ethical guidelines for educational research, fourth edition, 2018. URL https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2018.

[29] Malcolm Brown. Learning analytics: moving from concept to practice. *EDUCAUSE Learning Initiative*, 7, 2012.

[30] Poppy Brown. *The Invisible Problem?: Improving Students' Mental Health*. Higher Education Policy Institute, 2016.

[31] John T Cacioppo and Wendi L Gardner. Emotion. *Annual review of psychology*, 50 (1):191–214, 1999.

[32] David Camacho, Álvaro Ortigosa, Estrella Pulido, and María D R-Moreno. Ai techniques for monitoring student learning process. In Francisco Jose Garcia Penalvo, editor, *Advances in E-Learning: Experiences and Methodologies*, pages 149–172. IGI Global, 2008.

[33] Walter B Cannon. The james-lange theory of emotions: A critical examination and an alternative theory. *The American journal of psychology*, 39(1/4):106–124, 1927.

[34] Rich Caruana and Alexandru Niculescu-Mizil. An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on Machine learning*, pages 161–168, 2006.

[35] Center for Advanced Research On Learning. CAROL Learner MOOC Data Set. `https://datastage.stanford.edu`.

[36] Soumaya Chaffar and Claude Frasson. Predicting learners' emotional response in intelligent distance learning systems. In *FLAIRS Conference*, pages 383–388, 2006.

[37] Doug Clow. An overview of learning analytics. *Teaching in Higher Education*, 18(6): 683–695, 2013.

[38] Francesco Colace, Massimo De Santo, and Luca Greco. Safe: A sentiment analysis framework for e-learning. *International Journal of Emerging Technologies in Learning*, 9(6), 2014.

[39] Jo Coldwell, Annemieke Craig, T Paterson, and Jamie Mustard. Online students: Relationships between participation, demographics and academic performance. *Electronic journal of e-learning*, 6(1):19–30, 2008.

[40] CollegeAtlas.org. 41 facts about online students, 2014. URL `https://www.collegeatlas.org/41-surprising-facts-about-online-students.html`.

[41] Andrea Cornwall and Rachel Jewkes. What is participatory research? *Social science & medicine*, 41(12):1667–1676, 1995.

[42] Antonio R Damasio. Emotion in the perspective of an integrated nervous system. *Brain research reviews*, 26(2-3):83–86, 1998.

[43] Charles Darwin. The expression ofthe emotions in man and animals. *New York: Philosophical Library*, 1872.

[44] Shane Dawson, Dragan Gašević, George Siemens, and Srecko Joksimovic. Current state and future trends: A citation network analysis of the learning analytics field.

In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, LAK 2014, pages 231–240, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450326643. doi: 10.1145/2567574.2567585. URL `https://doi.org/10.1145/2567574.2567585`.

[45] O. de Vel, A. Anderson, M. Corney, and G. Mohay. Mining e-mail content for author identification forensics. *SIGMOD Rec.*, 30(4):55,Äì64, December 2001. ISSN 0163-5808. doi: 10.1145/604264.604272. URL `https://doi.org/10.1145/604264.604272`.

[46] Francesca D'Errico, Marinella Paciello, and Luca Cerniglia. When emotions enhance students,Äô engagement in e-learning processes. *Journal of e-Learning and Knowledge Society*, 12(4), 2016.

[47] Swantje Dettmers, Ulrich Trautwein, Oliver Lüdtke, Thomas Goetz, Anne C Frenzel, and Reinhard Pekrun. Students' emotions during homework in mathematics: Testing a theoretical model of antecedents and achievement outcomes. *Contemporary Educational Psychology*, 36(1):25–35, 2011.

[48] Tismy Devasia, TP Vinushree, and Vinayak Hegde. Prediction of students performance using educational data mining. In *2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)*, pages 91–95. IEEE, 2016.

[49] John DeweyJ. Interest and effort in education. *Boston : Houghton Mifflin Co. ; Cambridge : Riverside Press*, 1913.

[50] Beth Dietz-Uhler and Janet E Hurn. Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of interactive online learning*, 12(1): 17–26, 2013.

[51] Charles Docherty, Derek Hoy, Helena Topp, and Kathryn Trinder. elearning techniques supporting problem based learning in clinical simulation. *International Journal of medical informatics*, 74(7-8):527–533, 2005.

[52] Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2012.

[53] Anastasia Efklides and Simone Volet. Emotional experiences during learning: Multiple, situated and dynamic. *Learning and Instruction*, 15, 2005.

[54] Paul Ekman. Basic emotions. *Handbook of cognition and emotion*, 98(45-60):16, 1999.

[55] Paul Ekman and Wallace V Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2):124, 1971.

[56] Tanya Elias. Learning analytics. *Learning*, pages 1–22, 2011.

[57] Garry Falloon and Elaine Khoo. Exploring young students' talk in ipad-supported collaborative learning environments. *Computers & Education*, 77:13–28, 2014.

[58] Christiane Fellbaum. Wordnet: An electronic lexical database cambridge. *MA: MIT Press*, 1998.

[59] Dianne L Ferguson, Ginevra Ralph, and Gwen Meyer. *Designing personalized learning for every student*. ASCD, 2001.

[60] Rebecca Ferguson and Simon Buckingham Shum. Social learning analytics: five approaches. In *Proceedings of the 2nd international conference on learning analytics and knowledge*, pages 23–33, 2012.

[61] Rebecca Ferguson, Andrew Brasher, Doug Clow, Adam Cooper, Garron Hillaire, Jenna Mittelmeier, Bart Rienties, Thomas Ullmann, and Riina Vuorikari. Research

evidence on the use of learning analytics: Implications for education policy, 2016. URL http://oro.open.ac.uk/48173/. ISBN: 978-92-79-64441-2 ISSN: 1831-9424 Catalogue number: LF-NA-28294-EN-N.

[62] Jeremy D Finn and Donald A Rock. Academic success among students at risk for school failure. *Journal of applied psychology*, 82(2):221, 1997.

[63] Javier Gil Flores and Cristina Granado Alonso. Using focus groups in educational research: Exploring teachers' perspectives on educational change. *Evaluation review*, 1995.

[64] John Fritz. Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education*, 14(2):89–97, 2011.

[65] John Fritz. Using analytics at umbc: Encouraging student responsibility and identifying effective course designs. *EDUCAUSE Center for Applied Research*, pages 1–11, 2013.

[66] Fraide A Ganotice Jr, Jesus Alfonso D Datu, and Ronnel B King. Which emotional profiles exhibit the best learning outcomes? a person-centered analysis of students,Äô academic emotions. *School Psychology International*, 37(5):498–518, 2016.

[67] David Garcia and Frank Schweitzer. Emotions in product reviews–empirics and models. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pages 483–488. IEEE, 2011.

[68] Dragan Gasevic, Yi-Shan Tsai, Shane Dawson, and Abelardo Pardo. How do we start? an approach to learning analytics adoption in higher education. *International*

Journal of Information and Learning Technology, June 2019. ISSN 2056-4880. doi: 10.1108/IJILT-02-2019-0024.

[69] Helena Gómez-Adorno, Juan-Pablo Posadas-Duran, Germán Ríos-Toledo, Grigori Sidorov, and Gerardo Sierra. Stylometry-based approach for detecting writing style changes in literary texts. *Computación y Sistemas*, 22(1):47–53, 2018.

[70] Javier Gonzalez-Sanchez, Mustafa Baydogan, Maria Elena Chavez-Echeagaray, Robert K Atkinson, and Winslow Burleson. Affect measurement: A roadmap through approaches, technologies, and data analysis. In Myounghoon Jeon, editor, *Emotions and Affect in Human Factors and Human-Computer Interaction*, pages 255–288. Elsevier, 2017.

[71] Sumit Goswami, Sudeshna Sarkar, and Mayur Rustagi. Stylometric analysis of bloggers age and gender. In *Third international AAAI conference on weblogs and social media*, 2009.

[72] Carmen Gregori-Signes and Begoña Clavel-Arroitia. Analysing lexical density and lexical diversity in university students‚Äô written discourse. *Procedia-Social and Behavioral Sciences*, 198:546–556, 2015.

[73] M Grush. Monitoring the pace of student learning: Analytics at rio salado college. *Campus Technology*, 14, 2011.

[74] Andreas Haag, Silke Goronzy, Peter Schaich, and Jason Williams. Emotion recognition using bio-sensors: First steps towards an automatic system. In Elisabeth André, Laila Dybkjær, Wolfgang Minker, and Paul Heisterkamp, editors, *Affective Dialogue Systems*, pages 36–48, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg. ISBN 978-3-540-24842-2.

[75] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The WEKA data mining software: an update. *SIGKDD Explorations*, 11(1):10–18, 2009.

[76] Hanover Research. Learning analytics for student progress tracking, 11 2016. URL `https://www.imperial.edu/docs/research-planning/7932-learning-analytics-for-tracking-student-progress/file`.

[77] Vasileios Hatzivassiloglou and Kathleen R. McKeown. Predicting the semantic orientation of adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, ACL ,Äô98/EACL ,Äô98, page 174,Äì181, USA, 1997. Association for Computational Linguistics. doi: 10.3115/976909.979640. URL `https://doi.org/10.3115/976909.979640`.

[78] Ángel Hernández-García, Inés González-González, Ana Isabel Jiménez-Zarco, and Julián Chaparro-Peláez. Applying social learning analytics to message boards in online distance learning: A case study. *Computers in Human Behavior*, 47:68–80, 2015.

[79] Raquel Hijon and Ángel Velázquez. E-learning platforms analysis and development of students tracking functionality. In *EdMedia+ Innovate Learning*, pages 2823–2828. Association for the Advancement of Computing in Education (AACE), 2006.

[80] Graeme Hirst and Vanessa Wei Feng. Changes in style in authors with alzheimer's disease. *English Studies*, 93(3):357–370, 2012.

[81] G. Holmes, A. Donkin, and I. H. Witten. Weka: a machine learning workbench. In *Proceedings of ANZIIS '94 - Australian New Zealnd Intelligent Information Systems Conference*, pages 357–361, Nov 1994. doi: 10.1109/ANZIIS.1994.396988.

[82] Mushtaq Hussain, Wenhao Zhu, Wu Zhang, Syed Muhammad Raza Abidi, and Sadaqat Ali. Using machine learning to predict student difficulties from learning session data. *Artificial Intelligence Review*, 52(1):381–407, 2019.

[83] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[84] William James. Mind association what is an emotion. *Mind*, 9(34):188–205, 1884.

[85] Jeroen Janssen, Gijsbert Erkens, and Paul A Kirschner. Group awareness tools: It's what you do with it that matters. *Computers in human behavior*, 27(3):1046–1058, 2011.

[86] Ioana Jivet, Maren Scheffel, Marcus Specht, and Hendrik Drachsler. License to evaluate: Preparing learning analytics dashboards for educational practice. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450364003. doi: 10.1145/3170358.3170421. URL `https://doi.org/10.1145/3170358.3170421`.

[87] Hae Youn Joung and Ellen Yi-Luen Do. Tactile hand gesture recognition through haptic feedback for affective online communication. In Constantine Stephanidis, editor, *Universal Access in Human-Computer Interaction. Users Diversity*, pages 555–563, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg. ISBN 978-3-642-21663-3.

[88] James Joyce. Bayes theorem. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2019 edition, 2019.

[89] Hiroshi Kanayama and Tetsuya Nasukawa. Fully automatic lexicon expansion for domain-oriented sentiment analysis. In *Proceedings of the 2006 Conference on Em-*

pirical Methods in Natural Language Processing, EMNLP 2006, pages 355–363, USA, 2006. Association for Computational Linguistics. ISBN 1932432736.

[90] Mehmed Kantardzic. *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons, 2011.

[91] Edward Chao-Chun Kao, Chun-Chieh Liu, Ting-Hao Yang, Chang-Tai Hsieh, and Von-Wun Soo. Towards text-based emotion detection a survey and possible improvements. In *2009 International Conference on Information Management and Engineering*, pages 70–74. IEEE, 2009.

[92] Sahil P Karkhanis and Shweta S Dumbre. A study of application of data mining and analytics in education domain. *International Journal of Computer Applications*, 120 (22), 2015.

[93] Kostas Karpouzis, George Caridakis, Loic Kessous, Noam Amir, Amaryllis Raouzaiou, Lori Malatesta, and Stefanos Kollias. Modeling naturalistic affective states via facial, vocal, and bodily expressions recognition. In Thomas S. Huang, Anton Nijholt, Maja Pantic, and Alex Pentland, editors, *Artifical Intelligence for Human Computing*, pages 91–112, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. ISBN 978-3-540-72348-6.

[94] Zied Kechaou, Mohamed Ben Ammar, and Adel M Alimi. Improving e-learning with sentiment analysis of users' opinions. In *2011 IEEE global engineering education conference (EDUCON)*, pages 1032–1038. IEEE, 2011.

[95] Alistair Kennedy and Diana Inkpen. Sentiment classification of movie reviews using contextual valence shifters. *Computational intelligence*, 22(2):110–125, 2006.

[96] ChanMin Kim, Seung Won Park, and Joe Cozart. Affective and motivational factors

of learning in online mathematics courses. *British Journal of Educational Technology*, 45(1):171–185, 2014.

[97] Soo-Min Kim and Eduard Hovy. Determining the sentiment of opinions. In *Proceedings of the 20th International Conference on Computational Linguistics*, COLING 2004, pages 1367,s, USA, 2004. Association for Computational Linguistics. doi: 10.3115/1220355.1220555. URL `https://doi.org/10.3115/1220355.1220555`.

[98] Ronnel B King and Shaljan Areepattamannil. What students feel in school influences the strategies they use for learning: Academic emotions and cognitive/meta-cognitive strategies. *Journal of Pacific Rim Psychology*, 8(1):18–27, 2014.

[99] Ronnel B King and Marianne Jennifer M Gaerlan. High self-control predicts more positive emotions, better engagement, and higher achievement in school. *European journal of psychology of education*, 29(1):81–100, 2014.

[100] Joris Klerkx, Katrien Verbert, and Erik Duval. Enhancing learning with visualization techniques. In J. Michael Spector, M. David Merrill, Jan Elen, and M. J. Bishop, editors, *Handbook of Research on Educational Communications and Technology*, pages 791–807. Springer New York, New York, NY, 2014. ISBN 978-1-4614-3185-5. doi: 10.1007/978-1-4614-3185-5_64. URL `https://doi.org/10.1007/978-1-4614-3185-5_64`.

[101] Marileena Koskela, Piia Kiltti, Inka Vilpola, and Janne Tervonen. Suitability of a virtual learning environment for higher education. *Electronic Journal of e-Learning*, 3 (1):23–32, 2005.

[102] Yasuhiro Kotera, Elaine Conway, and William Van Gordon. Mental health of uk

university business students: Relationship with shame, motivation and self-compassion. *Journal of Education for Business*, 94(1):11–20, 2019.

[103] Vitomir Kovanović, Dragan Gašević, Shane Dawson, Srećko Joksimović, Ryan S. Baker, and Marek Hatala. Penetrating the black box of time-on-task estimation. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, LAK 2015, pages 184–193, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450334174. doi: 10.1145/2723576.2723623. URL `https://doi.org/10.1145/2723576.2723623`.

[104] Uros Krcadinac, Philippe Pasquier, Jelena Jovanovic, and Vladan Devedzic. Synesketch: An open source library for sentence-based emotion recognition. *IEEE Transactions on Affective Computing*, 4(3):312–325, 2013.

[105] Uros Krcadinac, Jelena Jovanovic, Vladan Devedzic, and Philippe Pasquier. Textual affect communication and evocation using abstract generative visuals. *IEEE Transactions on Human-Machine Systems*, 46(3):370–379, 2015.

[106] LB Krithika. Student emotion recognition system (sers) for e-learning improvement based on learner concentration metric. *Procedia Computer Science*, 85:767–776, 2016.

[107] Andrew E. Krumm, R. Joseph Waddington, Stephanie D. Teasley, and Steven Lonn. A learning management system-based early warning system for academic advising in undergraduate engineering. In Johann Ari Larusson and Brandon White, editors, *Learning Analytics: From Research to Practice*, pages 103–119. Springer New York, New York, NY, 2014. ISBN 978-1-4614-3305-7. doi: 10.1007/978-1-4614-3305-7_6. URL `https://doi.org/10.1007/978-1-4614-3305-7_6`.

[108] Dana Kulic and Elizabeth A Croft. Affective state estimation for human–robot interaction. *IEEE Transactions on Robotics*, 23(5):991–1000, 2007.

[109] Nicholas Lester, Laurie Feldman, and Fermín Moscoso del Prado Martín. You can take a noun out of syntax...: Syntactic similarity effects in lexical priming. In Richard Granger, Ulrike Hahn, and Richard Sutton, editors, *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, COGSCI 2017, pages 2537–2542. London Computational Foundations of Cognition, 2017.

[110] David D. Lewis. Naive (bayes) at forty: The independence assumption in information retrieval. In Claire Nédellec and Céline Rouveirol, editors, *Machine Learning: ECML-98*, pages 4–15, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg. ISBN 978-3-540-69781-7.

[111] Fei Li, Jingyao Qi, Guiwei Wang, and Xiaofeng Wang. Traditional classroom vs e-learning in higher education: Difference between students' behavioral engagement. *International Journal of Emerging Technologies in Learning (iJET)*, 9(2):48–51, 2014.

[112] Diane J. Litman and Kate Forbes-Riley. Predicting student emotions in computer-human tutoring dialogues. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, ACL ,Äô04, page 351,Äìes, USA, 2004. Association for Computational Linguistics. doi: 10.3115/1218955.1219000. URL `https://doi.org/10.3115/1218955.1219000`.

[113] Bing Liu. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167, 2012.

[114] Shuai Liu, Lulu Yang, Chenxi Zhang, Yu-Tao Xiang, Zhongchun Liu, Shaohua Hu,

and Bin Zhang. Online mental health services in china during the covid-19 outbreak. *The Lancet Psychiatry*, 7(4):e17–e18, 2020.

[115] Hugo Lövheim. A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical hypotheses*, 78(2):341–348, 2012.

[116] Peter Lunt and Sonia Livingstone. Focus groups in communication and media research. *Journal of Communication*, 42:78–87, 1996.

[117] Griet Lust, Mieke Vandewaetere, Eva Ceulemans, Jan Elen, and Geraldine Clarebout. Tool-use in a blended undergraduate course: In search of user profiles. *Computers & Education*, 57(3):2135–2144, 2011.

[118] Griet Lust, Jan Elen, and Geraldine Clarebout. Regulation of tool-use within a blended course: Student differences and performance effects. *Computers & Education*, 60(1):385–395, 2013.

[119] Wincenty Lutoslawski. Principes de stylométrie appliqués à la chronologie des œuvres de platon. *Revue des études grecques*, 11(41):61–81, 1898.

[120] Stephen G MacDonell, Donna Buckingham, Andrew R Gray, and Philip J Sallis. Software forensics: extending authorship analysis techniques to computer programs. *JL & Inf. Sci.*, 13:34, 2002.

[121] Dana-Kristin Mah. Learning analytics and digital badges: Potential impact on student retention in higher education. *Technology, Knowledge and Learning*, 21(3):285–305, 2016.

[122] Isa Maks and Piek Vossen. A lexicon model for deep sentiment analysis and opinion mining applications. *Decision Support Systems*, 53(4):680–688, 2012.

[123] Mika V Mäntylä, Daniel Graziotin, and Miikka Kuutila. The evolution of sentiment analysis,Äîa review of research topics, venues, and top cited papers. *Computer Science Review*, 27:16–32, 2018.

[124] Farshid Marbouti, Heidi A Diefes-Dux, and Krishna Madhavan. Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education*, 103:1–15, 2016.

[125] Colin Martindale and Dean McKenzie. On the utility of content analysis in author attribution: The federalist. *Computers and the Humanities*, 29(4):259–270, 1995.

[126] Karen D Mattingly, Margaret C Rice, and Zane L Berge. Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning: An International Journal*, 4(3):236–247, 2012.

[127] Merry McDonald, Brian Dorn, and Gary McDonald. A statistical analysis of student performance in online computer science courses. In *Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education*, SIGCSE ,Äô04, page 71,Äì74, New York, NY, USA, 2004. Association for Computing Machinery. ISBN 1581137982. doi: 10.1145/971300.971327. URL `https://doi.org/10.1145/971300.971327`.

[128] Judith L Meece, Phyllis C Blumenfeld, and Rick H Hoyle. Students' goal orientations and cognitive engagement in classroom activities. *Journal of educational psychology*, 80(4):514, 1988.

[129] Carolina Mega, Lucia Ronconi, and Rossana De Beni. What makes a good student? how emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of educational psychology*, 106(1):121, 2014.

[130] Jianyu Miao and Lingfeng Niu. A survey on feature selection. *Procedia Computer Science*, 91:919–926, 2016.

[131] George A Miller. *WordNet: An electronic lexical database*. MIT press, 1998.

[132] Saif M. Mohammad and Peter D. Turney. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, CAAGET ,Äô10, page 26,Äì34, USA, 2010. Association for Computational Linguistics.

[133] Atasi Mohanty. Affective pedagogical agent in e-learning environment: A reflective analysis. *Creative Education*, 7(04):586, 2016.

[134] Jennifer Moon. Learning journals and logs, reflective diaries. *Centre for teaching and learning good practice in teaching and learning*, pages 1–29, 2003.

[135] David L Morgan. Focus group as qualitative research. *Newbury Park, CA: Sage Publications Inc.*, 1988.

[136] David L Morgan. Focus groups. *Annual review of sociology*, 22(1):129–152, 1996.

[137] David L Morgan, Richard A Krueger, and Jean A King. *The focus group kit, Vols. 1–6*. Sage Publications, Inc, 1998.

[138] Frederick Mosteller and David L. Wallace. *Applied Bayesian and Classical Inference: The Case of The Federalist Papers*, chapter 1 - The Federalist Papers As a Case Study, pages 1–15. Springer New York, New York, NY, 1984. ISBN 978-1-4612-5256-6. doi: 10.1007/978-1-4612-5256-6_1. URL `https://doi.org/10.1007/978-1-4612-5256-6_1`.

[139] Youssef Mourdi, Mohamed Sadgal, Hamada El Kabtane, and Wafaa Berrada Fathi. A machine learning-based methodology to predict learners' dropout, success or failure in moocs. *International Journal of Web Information Systems*, 2019.

[140] Ahmed Mueen, Bassam Zafar, and Umar Manzoor. Modeling and predicting students' academic performance using data mining techniques. *International Journal of Modern Education and Computer Science*, 8(11):36, 2016.

[141] Krista R Muis, Cynthia Psaradellis, Susanne P Lajoie, Ivana Di Leo, and Marianne Chevrier. The role of epistemic emotions in mathematics problem solving. *Contemporary Educational Psychology*, 42:172–185, 2015.

[142] Myriam Munezero, Calkin Suero Montero, Maxim Mozgovoy, and Erkki Sutinen. Exploiting sentiment analysis to track emotions in students‚Äô learning diaries. In *Proceedings of the 13th Koli Calling International Conference on Computing Education Research*, Koli Calling ‚Äô13, page 145‚Äì152, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450324823. doi: 10.1145/2526968.2526984. URL `https://doi.org/10.1145/2526968.2526984`.

[143] Myriam D Munezero, Calkin Suero Montero, Erkki Sutinen, and John Pajunen. Are they different? affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE transactions on affective computing*, 5(2):101–111, 2014.

[144] David G Myers. Theories of emotion. *Psychology: Seventh Edition, New York, NY: Worth Publishers*, 500, 2004.

[145] Tobias O. Nyumba, Kerrie Wilson, Christina J Derrick, and Nibedita Mukherjee. The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and evolution*, 9(1):20–32, 2018.

[146] Felicia K Oluwalola. Effect of emotion on distance e-learning–the fear of technology. *International Journal of Social Science and Humanity*, 5(11):966, 2015.

[147] Alvaro Ortigosa, José M Martín, and Rosa M Carro. Sentiment analysis in facebook and its application to e-learning. *Computers in human behavior*, 31:527–541, 2014.

[148] Andrew Ortony, Donald A Norman, and William Revelle. Affect and proto-affect in effective functioning. *Who needs emotions*, pages 173–202, 2005.

[149] Bo Pang and Lillian Lee. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, ACL 2004, page 271es, USA, 2004. Association for Computational Linguistics. doi: 10.3115/1218955.1218990. URL https://doi.org/10.3115/1218955.1218990.

[150] Zachary A Pardos, Anthony Whyte, and Kevin Kao. moocrp: Enabling open learning analytics with an open source platform for data distribution, analysis, and visualization. *Technology, Knowledge and Learning*, 21(1):75–98, 2016.

[151] Andrew Parker and Jonathan Tritter. Focus group method and methodology: current practice and recent debate. *International Journal of Research & Method in Education*, 29(1):23–37, 2006.

[152] W Gerrod Parrott. *Emotions in social psychology: Essential readings*. Psychology Press, 2001.

[153] Reinhard Pekrun. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*, 18(4):315–341, 2006.

[154] Reinhard Pekrun and Lisa Linnenbrink-Garcia. Academic emotions and student engagement. In Sandra L. Christenson, Amy L. Reschly, and Cathy Wylie, editors, *Handbook of Research on Student Engagement*, pages 259–282. Springer US, Boston, MA, 2012. ISBN 978-1-4614-2018-7. doi: 10.1007/978-1-4614-2018-7_12. URL https://doi.org/10.1007/978-1-4614-2018-7_12.

[155] Reinhard Pekrun, Thomas Goetz, Wolfram Titz, and Raymond P Perry. Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist*, 37(2):91–105, 2002.

[156] Reinhard Pekrun, Thomas Goetz, Anne C Frenzel, Petra Barchfeld, and Raymond P Perry. Measuring emotions in students,Äô learning and performance: The achievement emotions questionnaire (aeq). *Contemporary educational psychology*, 36(1):36–48, 2011.

[157] Reinhard Pekrun, Nathan C Hall, Thomas Goetz, and Raymond P Perry. Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology*, 106(3):696, 2014.

[158] Caroline A Petrarca, Julia Warner, Andrew Simpson, Robert Petrarca, Abdel Douiri, David Byrne, and Timothy L Jackson. Evaluation of elearning for the teaching of undergraduate ophthalmology at medical school: a randomised controlled crossover study. *Eye*, 32(9):1498, 2018.

[159] Gabriele Piccoli, Rami Ahmad, and Blake Ives. Web-based virtual learning environments: A research framework and a preliminary assessment of effectiveness in basic it skills training. *MIS quarterly*, pages 401–426, 2001.

[160] Matthew D Pistilli, Kimberly Arnold, and Matt Bethune. Signals: Using academic analytics to promote student success. *EDUCAUSE Review Online*, 10(8), 2012.

[161] Robert Plutchik. A general psychoevolutionary theory of emotion. In Robert Plutchik and Henry Kellerman, editors, *Theories of emotion*, pages 3–33. Academic Press, 1980.

[162] Soujanya Poria, Erik Cambria, Newton Howard, Guang-Bin Huang, and Amir Hussain. Fusing audio, visual and textual clues for sentiment analysis from multimodal content. *Neurocomputing*, 174:50–59, 2016.

[163] Fatemeh Rabiee. Focus-group interview and data analysis. *Proceedings of the nutrition society*, 63(4):655–660, 2004.

[164] Jacqueleen A Reyes. The skinny on big data in education: Learning analytics simplified. *TechTrends*, 59(2):75–80, 2015.

[165] Bart Rienties, Avinash Boroowa, Simon Cross, Chris Kubiak, Kevin Mayles, and Sam Murphy. Analytics4action evaluation framework: A review of evidence-based learning analytics interventions at the open university uk. *Journal of Interactive Media in Education*, 2016(1), 2016.

[166] Jennifer L. Robison, Scott W. Mcquiggan, and James C. Lester. Modeling task-based vs. affect-based feedback behavior in pedagogical agents: An inductive approach. In *Proceedings of the 2009 Conference on Artificial Intelligence in Education: Building Learning Systems That Care: From Knowledge Representation to Affective Modelling*, page 25,Äì32, NLD, 2009. IOS Press. ISBN 9781607500285.

[167] Juan D Rodriguez, Aritz Perez, and Jose A Lozano. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE transactions on pattern analysis and machine intelligence*, 32(3):569–575, 2009.

[168] Stephen Rollnick and William R Miller. What is motivational interviewing? *Behavioural and cognitive Psychotherapy*, 23(4):325–334, 1995.

[169] VL Rubin, JM Stanton, and ED Liddy. Discerning emotions in texts, the aaai symposium on exploring attitude and affect in text, aaai-eaat 2004, 2004.

[170] Penny Rush. Isolation and connection: the experience of distance education. *International Journal of E-Learning & Distance Education*, 30(2), 2015.

[171] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980.

[172] Yvan Saeys, Iñaki Inza, and Pedro Larrañaga. A review of feature selection techniques in bioinformatics. *bioinformatics*, 23(19):2507–2517, 2007.

[173] Pradeep Sahu. Closure of universities due to coronavirus disease 2019 (covid-19): impact on education and mental health of students and academic staff. *Cureus*, 12(4), 2020.

[174] J Fernando Sánchez-Rada and Carlos A Iglesias. Onyx: A linked data approach to emotion representation. *Information Processing & Management*, 52(1):99–114, 2016.

[175] Maren Scheffel, Hendrik Drachsler, Slavi Stoyanov, and Marcus Specht. Quality indicators for learning analytics. *Journal of Educational Technology & Society*, 17(4): 117–132, 2014.

[176] Klaus R Scherer and Harald G Wallbott. Evidence for universality and cultural variation of differential emotion response patterning. *Journal of personality and social psychology*, 66(2):310, 1994.

[177] B Schmarzo. What universities can learn from big data higher education analytics, 2014.

[178] Vanessa Scholes. The ethics of using learning analytics to categorize students on risk. *Educational Technology Research and Development*, 64(5):939–955, 2016.

[179] Niall Sclater, Alice Peasgood, and Joel Mullan. Learning analytics in higher education. *London: Jisc. Accessed February*, 8:2017, 2016.

[180] Philip Sedgwick. Pearson‚Äôs correlation coefficient. *Bmj*, 345:e4483, 2012.

[181] Nilesh M. Shelke, Shriniwas Deshpande, and Vilas Thakre. Approaches towards emotion extraction from text. *IJCA Proceedings on National Conference on Innovative Paradigms in Engineering & Technology 2013*, NCIPET 2013(4):10–14, 12 2013. Full text available.

[182] Shiv Naresh Shivhare and Saritha Khethawat. Emotion detection from text. *arXiv preprint arXiv:1205.4944*, 2012.

[183] George Siemens and Phil Long. Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46(5):30, 2011.

[184] Sabina Sisovic, Maja Matetic, and Marija Brkic Bakaric. Mining student data to assess the impact of moodle activities and prior knowledge on programming course success. In *Proceedings of the 16th International Conference on Computer Systems and Technologies*, CompSysTech ‚Äô15, page 366‚Äì373, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450333573. doi: 10.1145/2812428.2812459. URL `https://doi.org/10.1145/2812428.2812459`.

[185] Ellen A Skinner and Michael J Belmont. Motivation in the classroom: Reciprocal

effects of teacher behavior and student engagement across the school year. *Journal of educational psychology*, 85(4):571, 1993.

[186] Sharon Slade and Paul Prinsloo. Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10):1510–1529, 2013.

[187] Michael Small. Mba student perspectives on online degrees and courses: An empirical analysis. *International Journal on E-Learning*, 2(2):5–15, 2003.

[188] Craig A Smith and Phoebe C Ellsworth. Patterns of cognitive appraisal in emotion. *Journal of personality and social psychology*, 48(4):813, 1985.

[189] Dan Song, Hongfei Lin, and Zhihao Yang. Opinion mining in e-learning system. In *2007 IFIP international conference on network and parallel computing workshops (NPC 2007)*, pages 788–792. IEEE, 2007.

[190] PS Sreeja and GS Mahalakshmi. Emotion models: a review. *Int J Control Theory Appl*, 10:651–657, 2017.

[191] Pankaj Srivastava. Advantages & disadvantages of e-education & e-learning. *Journal of Retail Marketing & Distribution Management*, 2(3):22–27, 2019.

[192] Elizabeth Stamp, Lee Crust, Christian Swann, John Perry, Peter Clough, and David Marchant. Relationships between mental toughness and psychological wellbeing in undergraduate students. *Personality and Individual Differences*, 75:170–174, 2015.

[193] Kim Storrie, Kathy Ahern, and Anthony Tuckett. A systematic review: students with mental health problems,Äîa growing problem. *International journal of nursing practice*, 16(1):1–6, 2010.

[194] Carlo Strapparava and Rada Mihalcea. Learning to identify emotions in text. In *Proceedings of the 2008 ACM Symposium on Applied Computing*, SAC ,Äô08, page 1556,Äì1560, New York, NY, USA, 2008. Association for Computing Machinery. ISBN 9781595937537. doi: 10.1145/1363686.1364052. URL `https://doi.org/10.1145/1363686.1364052`.

[195] Priyavahani Subramanian, Nursyafeeka Zainuddin, Salem Alatawi, Tahmineh Javabdeh, and A Hussin. A study of comparison between moodle and blackboard based on case studies for better lms. *Journal of Information Systems Research and Innovation*, 6:26–33, 2014.

[196] Jong Hwan Suh. Comparing writing style feature-based classification methods for estimating user reputations in social media. *SpringerPlus*, 5(1):261, 2016.

[197] Jianwen Sun, Zongkai Yang, Sanya Liu, and Pei Wang. Applying stylometric analysis techniques to counter anonymity in cyberspace. *Journal of Networks*, 7(2):259, 2012.

[198] Hamid Tarmazdi, Rebecca Vivian, Claudia Szabo, Katrina Falkner, and Nickolas Falkner. Using learning analytics to visualise computer science teamwork. In *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE ,Äô15, page 165,Äì170, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450334402. doi: 10.1145/2729094.2742613. URL `https://doi.org/10.1145/2729094.2742613`.

[199] Simhachalam Thamarana. Role of e-learning and virtual learning environment in english language learning. *Teaching English Language and Literature: Innovative Methods and Practices, ELTAI Tirupati*, pages 61–62, 2016.

[200] C Thorley. Not By Degrees: Improving student mental health in the UK's Uni-

versities, Institute for Public Policy Research, 2017. URL `https://www.ippr.org/publications/not-by-degrees`.

[201] Silvan S Tomkins. *Affect imagery consciousness: Volume I: The positive affects*, volume 1. Springer publishing company, 1962.

[202] Vicki Trowler. Student engagement literature review. *The higher education academy*, 11(1):1–15, 2010.

[203] Peter D. Turney. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL ,Äô02, page 417,Äì424, USA, 2002. Association for Computational Linguistics. doi: 10.3115/1073083.1073153. URL `https://doi.org/10.3115/1073083.1073153`.

[204] United Nations Educational, Scientific and Cultural Organization (UNESCO). Beijing consensus on artificial intelligence and education, 2019. URL `https://unesdoc.unesco.org/ark:/48223/pf0000368303`.

[205] UK Universities. Student mental wellbeing in higher education, 2015.

[206] Universities UK. Stepchange: mentally healthy universities, 2020.

[207] Iris Vessey. Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences*, 22(2):219–240, 1991.

[208] Felicidad T Villavicencio and Allan BI Bernardo. Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology*, 83(2):329–340, 2013.

[209] Chin-Yeh Wang, Gwo-Dong Chen, Chen-Chung Liu, and Baw-Jhiune Liu. Design an empathic virtual human to encourage and persuade learners in e-learning systems. In *Proceedings of the First ACM International Workshop on Multimedia Technologies for Distance Learning*, MTDL ,Äô09, page 27,Äì32, New York, NY, USA, 2009. Association for Computing Machinery. ISBN 9781605587578. doi: 10.1145/1631111.1631117. URL `https://doi.org/10.1145/1631111.1631117`.

[210] Yaqing Wang, Quanming Yao, James T Kwok, and Lionel M Ni. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys (CSUR)*, 53 (3):1–34, 2020.

[211] Webindia123. Us-based blackboard inc enters indian educational market, 09 2017. URL `https://career.webindia123.com/career/education-news-events/2017/us-based-blackboard-inc-enters-indian-educational-market-10-09-2017.htm`.

[212] Merriam Webster. Merriam webster online, 2014.

[213] Martin Weller. *Virtual learning environments: Using, choosing and developing your VLE*. Routledge, 2007.

[214] Cynthia Whissell. Traditional and emotional stylometric analysis of the songs of Beatles Paul McCartney and John Lennon. *Computers and the Humanities*, 30(3): 257–265, 1996.

[215] Cynthia Whissell. Using the revised dictionary of affect in language to quantify the emotional undertones of samples of natural language. *Psychological reports*, 105(2): 509–521, 2009.

[216] Cynthia M Whissell. The dictionary of affect in language. In Robert Plutchik and Henry Kellerman, editors, *The measurement of emotions*, pages 113–131. Elsevier, 1989.

[217] Allan Wigfield and John T Guthrie. Engagement and motivation in reading. In Michael L Kamil, P David Pearson, Elizabeth Birr Moje, and Peter Afflerbach, editors, *Handbook of reading research*, volume 3, pages 403–422. Routledge, 2011.

[218] Alyssa Friend Wise. Designing pedagogical interventions to support student use of learning analytics. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, LAK ,Äô14, page 203,Äì211, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450326643. doi: 10.1145/2567574.2567588. URL `https://doi.org/10.1145/2567574.2567588`.

[219] Billy Tak Ming Wong. Learning analytics in higher education: An analysis of case studies. *Asian Association of Open Universities Journal*, 12(1):21–40, 2017.

[220] Sherry Woosley and Darlena Jones. The foundations of map-works. *Educational Benchmarking (EBI)*, 2011.

[221] Himank Yadav and Juliang Li. Social media writing style fingerprint. *arXiv preprint arXiv:1712.04762*, 2017.

[222] Mohamed Yassine and Hazem Hajj. A framework for emotion mining from text in online social networks. In *2010 IEEE International Conference on Data Mining Workshops*, pages 1136–1142. IEEE, 2010.

[223] Han Yu, Zhiqi Shen, Qiong Wu, and Chunyan Miao. Designing socially intelligent virtual companions. *arXiv preprint arXiv:1411.7090*, 2014.

[224] G Udny Yule. *The statistical study of literary vocabulary*. Cambridge: Cambridge University Press, 1944.

[225] George Udny Yule. On sentence-length as a statistical characteristic of style in prose: With application to two cases of disputed authorship. *Biometrica*, 30:363–390, 1939.

[226] Hangjin Zhang, Kevin Almeroth, Allan Knight, Monica Bulger, and Richard Mayer. Moodog: Tracking students' online learning activities. In *EdMedia+ Innovate Learning*, pages 4415–4422. Association for the Advancement of Computing in Education (AACE), 2007.

[227] Zheng Zhao, Fred Morstatter, Shashvata Sharma, Salem Alelyani, Aneeth Anand, and Huan Liu. Advancing feature selection research. *ASU feature selection repository*, pages 1–28, 2010.

[228] Rong Zheng, Jiexun Li, Hsinchun Chen, and Zan Huang. A framework for authorship identification of online messages: Writing-style features and classification techniques. *Journal of the American society for information science and technology*, 57(3):378–393, 2006.